

Läkartidningen

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN

ebm

2000-2002



I N T R O D U K T I O N
T I L L
E B M

ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Evidensbaserad medicin

Ett konkret redskap för vårdens verkstadsgolv

Se även sidorna 2722 och framåt samt under korrespondens

I detta nummer av Läkartidningen beskriver Mats Eliasson hur den systematiska litteraturöversikten utvecklats till ett instrument för att definiera aktuellt kunskapsläge inom medicinen.

Genom de insatser som gjorts av Cynthia Mulrow, Tom Chalmers och David Sackett – från början starkt understödda av American College of Physicians med tidskriften *Annals of Internal Medicine* – har kravet på exakthet i kliniska översikter vuxit sig starkt, och så småningom kommit att sammanfattas i begreppet evidensbaserad medicin

(EBM), en anglicism som slagit rot i svenskan.

Uttrycket evidensbaserad sjukvård har en klar praktisk inställning [1], i detta fall med ett angivet mål snarare än med en strikt definition av begreppet. Målet skulle vara att eliminera beslutsformer som leder till dyra, felaktiga och eventuellt farliga beslut [2].

EBM är sålunda en praktisk filosofi, som innebär att varje åtgärd i sjukvården skall vara baserad på bästa möjliga tillgängliga kunskap. Det innebär att läkaren när han ger en rekommendation skall veta hur väl underbyggd just denna rekommendation är. I bästa fall finns det en eller flera randomiserade studier vars resultat kan utnyttjas för ett ställningstagande, men i många fall saknas sådana.

Det kan emellertid finnas annan, nästan lika bra evidens i den vetenskapliga litteraturen, vilken han då bör kän-

na till. Saknas även sådana finns det i alla fall kunskap med svagare underlag, allt enligt en fallande skala som definierats något olika i olika sammanhang. Den är emellertid alltid baserad på de kriterier som först angavs av den kanadensiska arbetsgruppen [3], sedan den amerikanska, för förebyggande åtgärder och som tillämpats i flera av SBU:s (Statens beredning för medicinsk utvärdering) rapporter.

I värsta fall får man förlita sig på en gemensam uppfattning om vad som är det bästa sättet att lösa det föreliggande problemet [4]. Vad det gäller är alltså att behärska det aktuella kunskapsunderlaget för den fråga, det problem, man är i begrepp att ta ställning till.

Aktuell kunskap

Ett nyckelord i sammanhanget är »aktuella». Detta innebär inte bara att dåliga översikter inte kan användas utan

Författare

LARS WERKÖ

professor, Statens beredning för medicinsk utvärdering, Stockholm.

Evidensbaserad medicin i Läkartidningen

Från och med detta nummer och återkommande en gång per månad kommer Läkartidningen att publicera artiklar i och om evidensbaserad medicin (EBM). Det är en av de största enskilda satsningarna som Läkartidningen har gjort och anledningarna till detta är flera. Den för våra läsare kanske viktigaste är att EBM och den systematiska litteraturöversikten kan användas av alla för att ge underlaget för de kliniska besluten – den beprövade erfarenheten – ett stabilt vetenskapligt fundament.

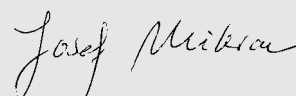
Som Lars Werkö redogör för i sin medicinska kommentar är principerna för EBM inte nya – men tillämpningen av en systematisk och enhetlig granskningsprocess har utmynnat i ett både kraftfullt och praktiskt använd-

bart hjälpmedel för välgrundade kliniska beslut, till nytta för såväl den enskilde patienten som patientgruppen och vårdorganisationen.

Flera av de stora internationella tidskrifterna, *British Medical Journal*, *JAMA*, *Lancet*, *New England Journal of Medicine* och andra, är sedan länge förespråkare för EBM. I detta perspektiv är EBM ingen nyhet för den svenska läkarkåren. I och med Läkartidningens satsning ökar tillgången på evidensbaserade översikter granskade och skrivna ur ett svenskt perspektiv och därmed de kliniskt verksamma läkarnas möjligheter att tillgodogöra sig denna viktiga kunskapskälla.

Statens beredning för medicinsk utvärdering, SBU, har länge arbetat för att sprida och utveckla EBM i Sve-

rige, och Läkartidningens nye EBM-redaktör Mats Eliasson har värvats från SBU på grundval av sina meriter som kunnig och effektiv projektledare på. Mats Eliassons medarbetarskap innebär dock inte att det föreligger varken formella eller informella bindningar mellan Läkartidningen och SBU. Läkartidningens uppgift är att kritiskt granska och föra ut kritiskt granskat material. Denna kritiska och obundna redaktionella hållning gäller allt material även från andra obundna och opartiska organisationer.



Josef Milerad
medicinsk chefredaktör

också att texter i läroböcker eller uppslagsverk, liksom rester av anteckningar från grundutbildningen, sällan är relevanta. Att läroböcker innehåller föråldrad lärdom är ett av de avslöjanden som gjorts under de senaste decennierna [5]. Detta gäller i synnerhet effekter av behandling som kan vara både under- och övervärderad i sådana texter.

För att hålla kunskapsförrådet aktuellt är det därför nödvändigt att följa litteraturen, eventuellt även elektronisk sådan, i stället för att lita till läroböcker. Detta är att begära mycket av den läkare som står mitt uppe i rutinsjukvårdens grottekvarn. Därför finns det hjälpmedel för att underlätta detta, i form av både elektroniska nätverk och pappersmedier [1, 6]. Till den uppräknade av källor som Eliasson gör kan läggas de publikationer som kommer från den mycket aktiva engelska motsvarigheten till SBU. Den enskilde läkaren måste ändå rationalisera sitt umgänge med dessa medier för att kunna hinna med.

Hur bara både aktuell och rationell?

John Fry, som även i Sverige måste betraktas som en av den moderna primärvårdens fäder, brukade framhålla att den genomsnittlige primärvårdsläkaren såg ett relativt begränsat antal problem i sin dagliga gärning. Han hade en tabell som visade att förekomst av en sjukdom som exempelvis tyreoidcancer var sådan att det fordrades ett tjugotal primärvårdsläkares upptagningsområde för att ett fall skulle uppträda under ett år – eller lika många år för en läkare innan han fick möta en patient med denna diagnos. Däremot såg den enskilde primärvårdsläkaren årligen flera hundra fall av övre luftvägsinfektion, många urinvägsinfektioner, något tiotal fall av kroppsliga symtom vid äktenskapliga kriser samt högst en hjärtinfarkt och kanske två slaganfall.

Detta innebär att den »vanlige» doktorn kan vara inställd på ett relativt litet antal problem som han ser ofta i sin



EBM är sålunda en praktisk filosofi, som innebär att varje åtgärd i sjukvården skall vara baserad på bästa möjliga tillgängliga kunskap. Det innebär att läkaren när han ger en rekommendation skall veta hur väl underbyggd just denna rekommendation är.

verksamhet. Hur många dessa problem är varierar med demografi och geografi, liksom vilka av dessa problem som är vanligast. Frånsett infektioner kan en sådan frågeställning vara behandling av måttligt högt blodtryck, en annan att ta hand om patienter med ont i ryggen. Avseende dessa, kanske ett tjugotal, gäller det att vara à jour med utvecklingen. Det innebär att läkaren regelbundet måste följa de databaser som avhandlar dessa problem. »Regelbundet» betyder kanske att en eller två gånger om året uppdatera kunskapen, vilket är helt möjligt om det är fråga om det tjugotal sådana återkommande problem man kan skilja ut.

Ställningstagande till den verkliga problempatienten, den som har en sjukdom som sällan dyker upp i rutinen, måste göras på ett annat sätt. Det är i sådana fall som det kan bli fråga om att själv definiera problemet, själv ställa upp ett frågeschema och ge sig ut på nätet eller söka andra databaser för att få de senaste kunskaperna [6]. I den genomsnittliga mottagningen borde detta bli fallet högst ett par gånger per år. Arbetar man på en vårdcentral med flera läkare kan sådana problem motivera ett personalmöte. Denna beskrivning av primärvårdens verksamhet kan tillämpas även på de stora kliniska specialiteterna.

Det är alltså inte fråga om att alltid

vara uppdaterad rörande allt som kan inträffa i den dagliga verksamheten. En anledning till kritik mot EBM, som man har ansett visa det orimliga i filosofin, har varit tron att EBM krävt aktuell uppdatering av all kunskap.

Har vi inte alltid arbetat med vetenskap som grund?

Kravet på evidensbaserad medicin får inte komma i motsatsställning till vad som alltid beskrivits som basen för all klinisk verksamhet: vetenskap och beprövad erfarenhet. Den evidensbaserade sjukvården är endast en vidareutveckling och mer preciserad definition av vad det innebär att grunda verksamheten på vetenskap. Den beprövade erfarenheten kommer in avseende det sätt på vilket den som är kliniskt verksam kan granska och utnyttja översikter, sammanställningar eller metaanalyser i förhållande till sin egen verksamhet.

Någon enstaka individ kanske besitter denna egenskap redan från början, men för flertalet är den något som måste övas upp. Den torde kunna utvecklas i en atmosfär av kritisk diskussion, under klinisk utbildning eller, varför inte, klinisk »fostran». Det är resultatet av denna process som utgör beprövad erfarenhet. Handedare, klinikchefer och hela vårdteamet har ett stort ansvar att se till att denna kritiska diskussion hålls levande.

Praktisk sjukvård versus klinisk forskning

Det kan också finnas anledning att betona att det är i praktisk sjukvård som EBM skall tillämpas. Den akademiska världen får förutsättas att alltid ha använt vetenskapliga kriterier på de litteratursammanställningar som ingår i forskningsuppgifter. Den kritiska inställning som många forskande specialister har till EBM kanske har sin grund i en känsla av hot från en dokumentalistbaserad verksamhet [7, 8]. Den är snarare ett uttryck för att man inte satt sig in i vad EBM egentligen innebär.

Det är framför allt den kliniska verksamheten – i öppen eller sluten vård – som har allt att vinna på EBM, inte den forskning som befinner sig i fronten.

Sätta sig in i begreppen

Många av dem som ger sig in i debatten för eller mot EBM gör det utan att ha satt sig in i vad begreppet egentligen står för [7, 8]. Till en del beror detta på att den rörelse som definierat begreppet och propagerat för det på olika sätt också är engagerad i andra parallellt förekommande rörelser i tiden [6]. Det centrum för begreppet EBM som finns i Oxford är starkt förknippat med Cochrane Collaboration, en annan stark institution, som är lokaliserad på samma ställe.

Det är dock viktigt att skilja på dem, då de står för helt olika filosofier. Cochrane Collaboration har tillkommit för att definiera, samla, granska och sprida resultat av randomiserade kliniska studier av åtgärder inom medicinen. Resultaten av dess verksamhet kan i stor utsträckning användas som underlag för de bevis som är av betydelse för stora delar av den evidensbaserade medicinen. De arbeten som Cochrane Collaboration publicerar i tidskrifter eller på nätet är således inte liktydiga med EBM, utan utgör en av många grundstenar för den.

Långt innan begreppet evidensbaserad medicin var uppfunnet varnade Alvan Feinstein för att bli alltför beroende av epidemiologers och statistikers uppfattning om kliniska syndrom och kliniska problem [9, 10]. Han var tidigt ute med att påpeka vikten av att låta den kliniska verkligheten beskrivas på ett sätt som skulle medge en vetenskaplig bearbetning av effekten av olika terapier av minst samma klass som den som använts i de stora multicenterstudierna [9]. Först på allra senaste tiden har de ledande börjat diskutera hur man skall kunna beskriva och bearbeta kvalitativa data på ett sådant sätt att de kan användas som underlag för jämförande studier.

Det är klinikern som har privilegiet att tolka resultaten av de stora multicenterundersökningarna avseende olika behandlingsmöjligheter. Det är endast han/hon som kan jämföra det patientmaterial som ingår i dessa undersökningar med den patient (eller de patienter) som han har ansvaret för i sin verksamhet. Detta har uttryckts bl a av John Swales, hypertoni-forskare och ansvarig för forskning i det brittiska National Health Service, när han konstaterar att det trots allt endast är den kliniskt verksamme som kan ge patienten den definitiva rekommendationen [11].

Är vårdprogram lösningen?

Införandet av vårdprogram, »guidelines», baserade på resultat av kontrollerade studier kan innebära ett framsteg just på grund av att de sammanställer vetandet på ett systematiskt sätt och ger råd om hur det skall tillämpas. Varje patient är emellertid unik, med sina biologiska och psykologiska förutsättningar och i sin relation till omvärlden. Resultaten av de stora studierna hänför sig till medeltal av stora grupper. För enhetlighetens skull utesluts ur dessa många patienter, kanske just de som har samma komplicerade bild som den patient jag just nu skall ta hand om. Vårdprogram måste därför tillämpas med omdöme, även om de baseras på resultat av kontrollerade studier.

Många specialister är inte särskilt positiva till vare sig vårdprogram eller begreppet EBM, eftersom de anser att de kan mer om både bakgrunden till de stora studierna och hur de skall tolkas. Nyligen har kardiologen John R Hampton tillsammans med ett antal engelska sjukvårdsforskare varnat för den byråkratiska kontroll som detaljerade vårdprogram kan leda till [12]. Det är visserligen införandet av NICE (National Institute of Clinical Excellence) som han är rädd för skall minska de engelska kardiologernas frihet att ta hand om sina patienter på bästa sätt. Förutsättningen för NICE är de vårdprogram som base-

ras på sådan litteraturgranskning som bl a Cochrane Collaboration står för.

Jag har full förståelse för de engelska kardiologernas rädsla. Det gäller emellertid att balansera mellan utnyttjande av bästa möjliga kunskap och de risker som ligger i att vara beroende av läkemedelsindustrins sponsring [13]. De främsta kardiologerna kan säkerligen slå vakt om sin frihet även från sponsorer. Granskning av fristående kritiker är dock nödvändig för att inte den stora massan av läkare skall bli förförda av resultat som baseras på oseriösa studier [13, 14].

Vetenskap är inte allt

Lika viktigt som att utnyttja den vetenskapliga kunskapen som underlag för sin verksamhet, lika viktigt är det att inte förlora den mänskliga kontakten med patienterna. Den alternativa medicinen popularitet sammanhänger – tyvärr – i viss utsträckning med att den vetenskapliga medicinens företrädare blivit alltför tekniska, alltför experimentellt inställda även i sin botande verksamhet. De kan därmed ha förlorat den nödvändiga känslan av empati.

Insikten om att det endast är en del av medicinen som kan ge verklig bot, att även den vetenskapliga kunskapsbasen är alltför begränsad i förhållande till de behov som föreligger, borde göra företrädarna för skolmedicinen mer ödmjuka, mer benägna att möta patienterna på deras nivå än vad som förekommer i många arroganta möten mellan själv-tillräckliga företrädare för vården – inte bara läkare – och lidande patienter.

Detta är så mycket viktigare som de förluster i trovärdighet som olämpligt professionellt uppträdande kan leda till lätt får alltför stort genomslag i allmänhetens föreställning om den medicinska vetenskapen överhuvudtaget. Oenighet inom olika grenar av medicinen bidrar lätt till en allmän tveksamhet i fråga om vad vetenskapen står för. Utvecklingen mot alternativ medicin kan ses som en del av den allmänna skepsis mot natur-

vetenskap som växer fram världen runt. Det finns till och med de som anser att denna vetenskap nått vägs ände.

Risken för att en sådan allmän negativ attityd sprider sig måste tas på allvar och motarbetas av dem som företräder forskning och vetenskap. Edith Efron påpekade vikten av detta för länge sedan när hon diskuterade cancerrisker [15], och New York Academy of Science har nyligen upprepat varningarna för att agera på ett sätt som skulle få allmänheten att förlora förtroendet för vetenskapen [16]. Rörelsen på väg mot evidensbaserad sjukvård, rätt avvägd, utgör det bästa sättet att framhålla den biomedicinska vetenskapens betydelse för en positiv utveckling av sjukvården.

Referenser

1. Levi R. Evidensbaserad sjukvård. Lund: Studentlitteratur, 1998.
2. Rosenberg W, Donald A. Evidence based medicine: an approach to problem-solving. *BMJ* 1995; 310: 1122-6.
3. Canadian task force on the periodic health examination: The periodic health examination. *Can Med Assoc J* 1979; 121: 1193-254.
4. Nordin-Johansson A, Asplund K. Randomised controlled trials and consensus as a basis for intervention in internal medicine. *J Intern Med* 2000; 247: 94-104.
5. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized controlled trials and recommendations of clinical experts. Treatment for myocardial infarction. *JAMA* 1992; 268: 240-8.
6. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. New York, Edinburgh, London: Churchill-Livingstone, 1997.
7. Vandebroucke JP. Evidence based medicine and »medecine d'observation». *J Clin Epidemiol* 1996; 49: 1335-8.
8. Tanenbaum SJ. What physicians know. *N Engl J Med* 1993; 329: 1268-9.
9. Feinstein AR. An additional basic science for clinical medicine. *Ann Intern Med* 1983; 99: 393-7, 544-50, 705-12.
10. Feinstein A. Two centuries of conflict-collaboration between medicine and mathematics. *J Clin Epidemiol* 1996; 49: 1339-43.
11. Swales J. The NHS research programme. *BMJ* 1997; 314: 466-7.
12. Miles A, Hampton JR, Hurwitz B, eds. NICE, CHI and the NHS Reforms. Enabling excellence or imposing control? London: Aesculapius Medical Press, 2000.
13. Johansen HK, Götsche PG. Problems in the design and reporting of trials of antifungal agents encountered during meta-analysis. *JAMA* 1999; 282: 1752-9.
14. Horton R. After Bezwoda. *Lancet* 2000; 355: 942-3.
15. Efron E. The apocalypics How environmental politic control what we know about cancer. New York, NY: Simon & Schuster, 1984.
16. Gross P, Levitt N, Lewis MW, eds. The flight from science and reason. *Ann NY Acad Sci* 1996; 775.

Att segla och navigera på medicinens hav

För inte mer än snart 20 år sedan mönstrade jag som ung AT-läkare på läkarskeppet. Kaptenen där, min första medicinöverläkare, visste vad som gällde. Kinidin vid förmaksflimmer och antiarytmika klass I vid frekventa ventrikulära extraslag. Även för övrigt var farlederna väl utmärkta.

Genom åren har jag blivit alltmer besvärad i närvaro av läkare som hela tiden vet hur det förhåller sig och hur saker och ting skall vara och behandlas. Min kliniska vardag visar nämligen föga likhet med den verklighet som dessa läkare tycks uppleva. Det är inte nog med att uppfattningar rent medicinskt kan gå isär och att det inte alltid är lätt att veta vilken behandlingsstrategi som numera gäller. Dessutom är mina patienter annorlunda än de medelpatienter som behandlingsrekommendationerna baserar sig på. »Panta rei», allt flyter, sa redan de gamla grekerna. Det är ingen dum idé att flyta med.

Men i längden går inte det heller särskilt bra. En utväg syntes mig vara sunda genomarbetade vårdprogram där klara rekommendationer utkristalliserats. Genom landstingets benägna bistånd fick jag möjlighet att delta i en kurs i Maastricht om kvalitetsutveckling av framför allt allmänläkarvård genom kvalitetsrevision med vårdprogram som bas. Men kan man verkligen jobba vårdprogrambaserat? Medan jag funderade drev jag in i en bokhandel och fann Sacketts bok »Evidence-based medicine» [1]. Där fann jag en verklighetsanpassad navigationshjälp.

Författarna inledde med att presentera en randomiserad studie över läkares fortbildning. Det enda som tycktes bita på läkarna och verkligen medföra en förändring i kliniskt beteende var kliniskt baserad EBM-metodik.

Väl hemkommen startade jag och några andra läkare en studiecirkel utifrån Sacketts bok. Var och en av delta-

garna i studiecirkeln fick i uppgift att ur sin kliniska vardag välja och formulera ett kliniskt problem och söka bästa tillgängliga evidens för att belysa frågeställningen. Det gällde sedan att se om en förändring i klinisk praxis var motiverad och tillämpbar i vår vårdssituation.

Förutom en träning i att söka i databaser innebar uppgiften en träning i att formulera frågor som går att besvara. Någon i gruppen ville undersöka CRP-reaktioner efter operativa ingrepp, en annan gruppledare ville studera antiviral terapi vid herpes zoster, en tredje diagnostik vid djup ventrombos, en fjärde om NSAID verkligen är bättre än paracetamol vid tendinit i armbåge, axel, skuldra och nacke, en femte vilka effekter kolinesterashämmare kan förväntas ha på senildementa patienter osv.

De studier vi fann granskades sedan kritiskt i förhållande till vår patientpopulation och våra nedskrivna frågeställningar. Särskilt fokuserades frågan om det effektmått som studien angav är det effektmått som är kliniskt relevant för patienten och för oss. Många studier föreföll märkvärdigt tunna redan vid denna granskning. Enkelt lärde vi oss översätta relativa riskreduktioner till absoluta riskreduktioner. Prislappen för behandlingsvinsten blev följesedeln.

Sedan då? Vad hände? Ja, mitt intryck efter kurserna är att det går att tillämpa evidensbaserad medicin även på en vårdcentral. Men arbetssättet behöver utvecklas och framför allt behöver forskningen anpassas till kliniskt relevanta frågeställningar som går att besvara. Att leta efter evidens som inte finns men som borde finnas blev för mig en av de första erfarenheterna av EBM.

En distriktsläkares vardag och behandlingsstrategier är oftast väl beprövade, inövade seglatser i kända vatten där navigeringsbehovet inte är så stort. Däremot är behovet av EBM-värderade diagnosmetoder överväldigande för en allmänmedicinare. Vilken diagnostisk kapacitet har detta prov i en population som min och med en förväntad prevalens av den sökta sjukdomen på x procent? Behovet av allmänmedicinsk forskning enbart kring dessa frågor förefaller stort.

EBM sparar läkares hjärna. En ele-

mentär utbildning i EBM lär oss att se vilken information och vilka uppgifter som är väsentliga i det koppel av information som biter efter oss varje dag i posten och från bildskärmen. Det som är tunt och det som är åsiktsmedicin och lösa påståenden kan gärna få dräneras förbi nervsystemet till återvinningen. Hjärnstressen minskar hos läkaren. Det blir då lättare att fokusera på den väsentliga informationen. En randomiserad, placebokontrollerad klinisk prövning på ett nytt läkemedel blir däremot svår att motstå. Vilken effekt kan jag rimligen uppnå med detta läkemedel? Vilken blir kostnaden? Utför dessa övningar t ex på Relenza och Vectavir. Numera frågar vi alltid läkemedelskonsulenterna om absoluta riskreduktioner i stället för relativa. Förvirringen hos dem blir ofta påtaglig. Är vi de enda läkare som ställer dessa frågor?

Sammanfattningsvis vill jag alltså mena att EBM är grundbulsten för en medicin som vill röra sig ute i verkligheten. Den är uppbyggd av miljontals, konsekvent utförda, bearbetade och kritiskt sammanställda observationer. Vilket annat sätt är möjligt om man tror på empiriskt grundad medicin? EBM står i kontrast till åsiktsmedicin, konsensusmedicin och vårdprogrammedicin. EBM ger motvikt till den ogrundade optimistmedicinen och för våra önskningar ner till det rimligas skala. Kanske inte alla vill se medicinerna i den skalstorleken?

Medicinens hav låter sig till slut ändå definieras. Även om strandlinjen kan variera är den ändå definierbar. Ett överlappande nät av observationer ger koordinater från vilka den fortsatta seglatsen kan fortgå. Går det åt icke förväntat håll får väl navigatören, som vid all orientering, noggrant iakttäta omgivningarna, reorientera sig och dra vidare. Under den seglatsen hoppas jag på förliga vindar.

Författare

GÖRAN WALLER
distriktsläkare, Grytnäs vårdcentral,
Kalix.

E-post: goran.waller@nll.se

Referenser

1. Sackett DL, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice & teach EBM. London: Churchill Livingstone, 1997.

Evidensbaserad medicin ”begins at home”

Incitament för bättre beslutsunderlag finns i den kliniska vardagen

I det dagliga mötet med patienter fattar läkare hundratals beslut: vid sjukhussängen, på akuten eller öppna mottagningar, på specialist- eller primärvårdsmottagningar, vid operationsbordet. Många andra kolleger befinner sig i samma pressade beslutssituation utan att direkt möta patienten: framför mikroskopet, vid granskning av röntgenbilden, elfloresen eller EKG-remsan. Det är ofta svåra val avseende fortsatt utredning, behandling eller vårdnivå.

Hur den kliniska beslutssituationen kan se ut i vardagen på en medicinavdelning exemplifieras i Ruta 1. Andra verksamheter kan exemplifiera liknande svåra val, som måste leda fram till snabba beslut för att det ska vara »flyt» i organisationen.

Svårt att alltid fatta rätt beslut

Vet vi då alltid vad det bästa valet är för varje enskild patient? Kan vi med gott samvete påstå att alla dessa vardagliga beslut baseras på »systematiska sammanställningar av bästa tillgängliga kunskap»? Om jag skärskådar min egen verksamhet, och granskar några frågor som kan väckas under en rond på en allmän internmedicinsk vårdavdelning (Ruta 2) agerar jag ofta utifrån en blandning av kunskapsfragment från fortbildningskurser, diskussioner med kolleger och rön baserade på traditionella (icke-systematiska) översikter i tidsskrifter.

Enstaka gånger kan det ha slumpat sig så väl att kommentarer från Läke-medelsverket eller Socialstyrelsen kan vara direkt kliniskt relevanta, dock sällan för den sortens frågor jag listat. Inte heller är Läke-medelsboken eller traditionella läroböcker till stor nytta. En

Författare

MATS ELIASSON

med dr, överläkare, Sunderby sjukhus, Luleå; projektsamordnare, Statens beredning för medicinsk utvärdering (SBU); medicinsk redaktör, Läkartidningen (evidensbaserad medicin).

E-post: Mats.Eliasson@nll.se

1: Exempel på vardagliga kliniska optioner som kräver snabba beslut

Utredning

- Är utredning meningsfull eller är observation på avdelning bästa insatsen?
- Ordinera undersökning med röntgen, på kemiskt laboratorium eller av konsulter.
- Remiss till paramedicinsk personal (sjukgymnast, arbetsterapeut, kurator).
- Samtal med anhöriga, fördjupad anamnes/status.
- Journalgenomgång (egna och andra kliniker); ring ordinarie läkare.

Behandling

- Är behandling överhuvudtaget meningsfull eller etiskt försvarbar?
- Skall jag börja med att sätta ut läkemedel? I så fall vilka?
- Kan mobilisering, vätska och nutrition vara tillräcklig behandling?
- Kan depression eller demens ligga bakom symtomen?
- Är nya läkemedel nödvändiga?

Vårdnivå

- Hör patienten överhuvudtaget hemma på min klinik?
- Skulle patienten ha mer nytta av rehabilitering eller strokeenhet?
- Vore det bästa att snarast skicka hem patienten och låta distriktsläkare/-sköterska göra hembesök?
- Är patienten »färdigbehandlad»?

SBU-rapport skulle kunna ge en tillförlitlig vägledning, men är sällan fokuserad på denna typ av frågeställningar.

Hur vi till slut väljer att agera är nog oftast baserat på en blandning av alla dessa källor, kryddad med en del »intuition». Är det i själva verket detta som är »beprövad erfarenhet»? Många av oss har varit med om att kandidaten eller underläkaren troskyldigt frågat varför vi valde att agera just så i detta fall, och upplevt frustrationen när vi skall försöka reda ut och verbalisera den rationella basen för beslutet.

Vi har i allför liten utsträckning bekräftat existensen av den osäkerhet som råder kring mängder av för oss – dock inte för patienten – vardagliga kliniska beslut. Varje läkare tvingas dagligen att ta ställning till ett stort antal frågor där svaren inte finns omedelbart tillgängliga och tiden saknas att söka dem. Efter-som verksamheten inte kan invänta ett bättre beslutsunderlag agerar vi så gott vi kan och går vidare. Att på detta sätt ständigt tvingas fatta viktiga beslut på

bristande underlag är frustrerande. Det bidrar till kårens känsla av att arbetet är tungt och svårt.

Evidensbaserad medicin som verktyg i klinisk vardag

Kan evidensbaserad medicin (EBM) förbättra situationen och leda till bättre kunskapsunderlag och därmed större (känsla av) kontroll över den kliniska vardagen? Jag är övertygad om att så är fallet, men inte utan visst arbete för den som vill utöva ett sådant arbetssätt.

För det första tar EBM sin utgångspunkt i kliniska och patientrelaterade frågeställningar, inte i det som forskarna råkar vara intresserade av. T ex ställer den EBM-orienterade klinikern frågan huruvida antileukotriener är bättre än långverkande betastimulerare för livskvaliteten hos patienter med astma, medan forskaren är mer intresserad av huruvida preparaten i större utsträckning än placebo minskar histamininducerad bronk obstruktion. EBM-doktorn är mer intresserad av frågan huruvida

bifosfonater minskar nya symtomgivande frakturer hos 75-åriga kvinnor än huruvida de har någon inverkan på ben-tätheten hos kvinnor som just blivit menopausala. Och gör bifosfonaterna i dessa situationer verkligen mer nytta än rejäla doser kalk och D-vitamin, än gång- och balansträning eller höft-skydd?

För det andra kan EBM på allt fler områden ställa lätt tillgängliga och kontinuerligt uppdaterade kunskapssammanställningar till klinikers förfogande. För det tredje kan EBM, i de fall systematiska sammanställningar saknas, ge de verktyg som behövs för att varje läkare skall få en nöjaktig uppfattning om svaret på sin kliniska fundering.

Denna arbetsprocess för ett evidensbaserat svar är krävande både att lära sig och att genomföra, men den är lika viktig som att lära sig nya metoder för behandling och diagnostik, och den kommer att få större genomslagskraft såväl i patientarbetet som för patientnyttan på längre sikt.

Evidensbaserad praktik är livslångt, självstyrt lärande

Evidensbaserad medicin är att fatta beslut om handläggning av patienter på basen av bästa tillgängliga kunskap. Det kan emellertid vara komplementärt att se på EBM som ett förhållningssätt eller som en arbetsmetod. Att praktisera EBM blir ett livslångt, självstyrt lärande där patientarbetet skapar ett behov av kliniskt viktig information om diagnos, prognos och terapi eller andra kliniska aspekter. Att se EBM som en arbetsprocess blir då en enkel utgångspunkt för lärandet (Ruta 3).

Resultatet kan givetvis bli att någon redan har utfört de senare tre stegen, och att ett färdigt kunskapsunderlag föreligger för oss att ta ställning till och agera utifrån. Underlaget kan vara en systematisk översikt från Cochranesamarbetet eller en medicinsk metodvärdering från SBU eller liknande organisationer. Under överskådlig framtid kommer dock inte sådana källor att kunna tillgodose vårt behov på många »smala» kliniska områden, men vi kan vänta oss alltmer stöd i övergripande frågor som rör stora patientgrupper, t ex: Är SSRI-preparat mer kostnadseffektiva än tricykliga preparat vid lindrig depression? För vilka patienter med grav obesitas är kirurgisk åtgärd ett bra val för att minska mortaliteten? Mår patienten med non-ulcusdyspepsi bättre av att *Helicobacter eradikeras*?

Konsten att prioritera

I det följande vill jag peka på ett möjligt sätt att arbeta med sådana vardagliga frågeställningar som inte lätt låter sig besvaras utifrån traditionella kunskaps-

2: Vardagliga frågor vid rond

- Skall jag ge antibiotikan intravenöst eller peroralt?
- Kan en demensutredning med åtföljande diagnos och eventuell behandling vara till nytta för patienten?
- Vill jag verkligen veta orsaken till anemin hos den mycket gamla och skröpliga patienten?
- Vilken insulinregim är bäst för patienten som sviktat på peroral medicinering?
- Skall jag behandla eller ytterligare utreda (och i så fall med vad?) patienten med B 12 i nedre området?
- Vinner patienten med viss progress i sin stroke något på att få heparin, men är inte riskerna större? Och skulle jag i så fall lika väl kunna använda lågmolekylära hepariner?
- Är det sannolikt att en akut datortomografi av hjärnan på den plötsligt djupt medvetslöse äldre patienten skall visa något som kan åtgärdas och vara av bestående nytta för patienten?
- Hur högt skall jag låta kreatinivärdet stiga hos den gravt sviktande patienten med ACE-hämmare där inga andra behandlingsalternativ kvarstår?
- Skall jag ge COX-2-hämmare i stället för NSAID bara för att patienten är över 75 år eller tidigare haft »magkatarr»?
- Behöver verkligen patienten sitt SSRI-preparat, eller vågar jag sätta ut det och se om yrseln och illamåendet blir bättre?
- Vad är tilläggs-effekten av bifosfonater hos patienten som fått kotkompression under pågående behandling med kalk, D-vitamin och östrogen? Skall jag då sätta ut östrogen? Vilken bifosfonat skall jag välja?

ILLUSTRATION: LASSE PERSSON

källor. Utgångspunkten är just de återkommande frågor som ronder eller mottagningar genererar (Ruta 2). Eftersom vi inte kan starta med att söka svaren på alla frågor måste vi dock prioritera.

Ett intressant exempel utgör det engelska »general practitioner»-systemet, där man under arbetsveckan löpande noterar sådana frågor. Läkargruppen på vårdcentralen går regelbundet tillsam-

mans igenom frågorna, prioriterar sedan *en* fråga, som får bilda utgångspunkt för en planerad utbildningsaktivitet baserad på en EBM-process. På detta sätt når man ett lärande som utgår från de egna kliniska problemen, inte från vad externa krafter såsom sjukhus-specialister eller läkemedelsindustri anser bör prioriteras. Ett sådant »från toppen-och-nedåt»-perspektiv är annars idag alltför dominerande inom karens utbildning.

Fokusera den kliniska frågeställningen!

Att gå från ett svåravgränsat kliniskt informationsbehov till en distinkt och fokuserad frågeställning har visat sig vara en nyckel till förståelse av den evidensbaserade medicinen. Metoden är tillika ett behändigt och lättanvänt vardagligt verktyg. Erfarenheter, såväl internationellt som i Sverige, visar att denna övning är viktig för att stödja läkarens utveckling mot ett evidensbaserat arbetssätt. Det är rentav så, att många kan klara sig med enbart en reflektion kring frågeställningen och hur denna bäst bör besvaras, och ändå uppnå en

3: Så utövas EBM

Att utöva EBM innebär att

- Omvandla informationsbehov till frågor som kan besvaras;
- Spåra den bästa kunskapen med maximal effektivitet;
- Utvärdera litteraturen kritiskt avseende validitet och klinisk nytta;
- Omsätta resultatet i din egen kliniska praxis.

4: De fyra stegen i den fokuserade kliniska frågeställningen

**Patienten:
Vem vill jag behandla?**

Utifrån din patient – beskriv viktiga karakteristika

Hos äldre patienter med hjärtsvikt till följd av dilaterad kardiomyopati som har sinusrytm ...

**Interventionen:
Med vad vill jag behandla?**

Vilken är den viktigaste interventionen jag överväger?

... kan tillägg av anti-koagulantia utöver standard hjärtsvikts-behandling ...

**Jämförelsen:
I stället för...**

Vilket behandlingsalternativ vill jag jämföra med?

... jämfört med enbart standardterapi ...

**Effektmått:
Varför vill jag behandla?**

Vad hoppas jag uppnå med behandlingen?

Vilken skada riskerar jag att åsamka patienten?

... leda till lägre morbiditet eller mortalitet från tromboembolism. Är det värt den ökande risken för blödning?

betydande kompetensökning vad gäller att värdera klinisk forskning och dess relevans för de egna patienterna.

Även om frågeställningarna mycket väl kan röra diagnostik, prognos eller andra aspekter har det visat sig enklast och mest givande att börja med att diskutera olika interventioner. Den kliniska frågeställningen kan omformuleras till en fyrdelad fråga (Ruta 4), varur successivt bilden framträder av hur de kliniska studier bör se ut som skulle kunna besvara frågan. Redan nu får vi alltså anledning att reflektera över viktiga element vid den fortsatta läsningen av kliniska studier. Den fokuserade frågan kommer att ligga till grund för och möjliggöra en kostnadseffektiv litteratursökning.

Vem är patienten?

Viktiga patientkarakteristika, såsom ålder och kön, måste vägas in. Vi bör noggrant fundera över om vi verkligen diskuterar en välaggränsad sjukdomsentitet där diagnostiska kriterier är allmänt accepterade och i överensstämmelse med klinisk praxis i Sverige. T ex är många studier av kronisk prostatit helt baserade på mikroskopi och odlingar, inte på den typiska sjukdomsbild som oftast ligger till grund i vardagens arbete. Det innebär att vi sannolikt behandlar andra patientgrupper än de som ingått i studier, och följaktligen att fynden från odlingsstudierna inte självklart är till nytta för oss och våra patienter. Liknande problem uppträder vid sinuitstudier som använt punktion för diagnostik, vilket sällar fram en speciell undergrupp vars resultat inte helt självklart låter sig överföras till de patienter som klinisk diagnostik identifierat.

Många sjukdomar utvecklas från lindriga tillstånd till svåra slutstadier, t ex demens och Parkinsons sjukdom.

Vi har då anledning att fundera över huruvida fynden från studier utförda i tidiga stadier av en sjukdom har relevans för vår patient, vars sjukdom kan vara betydligt längre gången. Våra patienter är oftast äldre, lider av flera sjukdomar och står på fler mediciner än de patienter som inkluderats i de kliniska studierna vi kommer att hitta vid en litteratursökning. Denna komorbiditet och polyfarmaci kan få oss att tvivla på effekten av att tillföra fler läkemedel eller andra behandlingar.

Ytterligare ett problem är att många studier är utförda i en högspecialiserad sjukvårdsmiljö där möjligheterna att optimera behandlingsresultaten är långt bättre än dem vi själva någonsin skulle kunna uppnå. Vi bör därför vara skeptiska mot interventioner som förväntas utföras i primärvård eller på länsdelssjukhus, men där all dokumentation utgår från Mayo-kliniken eller liknande inrättningar. Var därför noggrann med att notera om de kliniska studierna utförts i en vårdmiljö liknande din egen!

Glädjande nog går forskningen mot att dels exkludera så få patienter som möjligt, dels inkludera deltagare från många vårdenheter, för att därmed få ett så brett spektrum av patienter och vårdgivare som möjligt. På grund av de heterogena patientgrupperna blir resultatet ofta en viss utspädning av behandlingseffekterna, men vi har å andra sidan anledning att tro att denna konservativa modell bättre avspeglar patientnyttan i verkliga livet.

Slutligen bör man ändå fråga sig om forskning på patienter som skiljer sig mycket från mina egna patienter verkligen helt saknar värde för mig? Självklart inte, men man bör från början ha ställt kraven höga för att veta hur den ideala studien bör se ut för att vara av nytta just för mig. I praktiken får man

ställa sig frågan om min patient i så hög grad avviker från dem som ingått i de kliniska studierna att jag inte kan använda resultaten. Här bör man hellre fria än fälla, och man bör ha goda biologiska skäl för att helt lägga fynden åt sidan.

Vilken är interventionen?

De interventioner som vi är intresserade av är inte bara läkemedel eller kirurgiska ingrepp, och inte enbart sådant som läkare utför. Vilken är t ex effekten av de sjukgymnastiska eller arbetsterapeutiska insatserna? Finns det någon bevisad nytta med kuratorns eller psykologens samtalsbehandling eller kognitiva beteendeterapi? Med ortopedverkstadens ingripande av diabetikern med fotsår? Med dietistens gruppbehandling av överviktiga, logopedens behandling av dysfagi?

Vi kan till och med ställa oss mer strukturella frågor, som huruvida specialistvård verkligen är bättre än primärvård för min patient (eller vice versa), eller om hospis är bättre än traditionell vård. Allt detta kan ses som interventioner vilkas effekt går att utvärdera med hjälp av randomiserade och kontrollerade kliniska studier.

Vi behöver reflektera över hur interventionen skall beskrivas mest distinkt för att ha bäst chans att hitta just de studier som berör vår fråga och minska risken för störande »brus», dvs att vi hittar alltför många ovidkommande artiklar i vår fortsatta litteratursökning. Vid läkemedelsbehandling får vi besluta oss för om vi är intresserade av en grupp läkemedel eller ett specifikt preparat (effekt av statiner vid förhöjt kolesterol, eller effekt av pravastatin?). Vill vi veta något om den senaste triptanen mot migrän, eller räcker det att se på effekten av hela gruppen?

Detta hänger ihop med den intres-

santa frågan om s k klasseffekter: när skall vi anse att de biologiska effekterna är tillräckligt lika inom en grupp av läkemedel för att dra slutsatsen att preparaten bör vara likvärdiga att behandla med? Just nu är frågan aktuell avseende bl a protonpumpshämmare, ACE-hämmare, statiner, bifosfonater och sulfonylureapreparat.

Vilken dos och vilken behandlingstid är den i Sverige använda? Man bör vara observant på att många studier från andra länder använder andra doseringsregimer än vi är vana vid, varvid resultaten inte omedelbart låter sig överföras till vår kliniska vardag. Å andra sidan kan det ibland vara en nyttig påminnelse om att vi egentligen inte behandlar på basen av bästa tillgängliga evidens. Vi vet bl a att hjärtsviktpatienter i allmänhet får doser av ACE-hämmare som är betydligt lägre än de som ursprungligen visade sig förbättra överlevnad och livskvalitet.

Hur lång uppföljning vill jag ha för att känna mig övertygad om att effekten kvarstår och risken för ovanliga biverkningar är liten? För behandlingar som utgör bot av akuta tillstånd är detta inte ett problem, men många preparat ges för behandling under flera år trots att de studier som ligger till grund för terapin kan vara begränsade till sex månader eller ett år.

Andra interventioner än läkemedel bör beskrivas så avgränsat och så specifikt som möjligt (interferensbehandling på ländrygg, fenestrerat gips med klack, minilaparotomi, problembaserad och deltagarstyrd gruppbehandling).

En konst att göra korrekta jämförelser

Alla randomiserade kliniska studier är kontrollerade, dvs det finns en kontrollgrupp för att kunna jämföra effekten av den nya behandlingen med något annat. Kontrollgruppens behandling har traditionellt oftast varit någon form av placebo, vilket i de flesta fall är möjligt att använda även i icke-farmakologiska studier. Det är nödvändigt och rimligt att nya preparat i första hand visar klinisk effekt utöver den rena placeboeffekten (som i studier av analgetika kan vara så stor som 30 procent). För klinikern är det förstås tryggt att veta att preparatet har klinisk effekt, men denna information är sällan tillräcklig när det gäller de vardagliga besluten där vi står inför valet mellan det nya preparatet (som är bättre än placebo) och den gamla behandlingen (som förhoppningsvis också har visats ha verkan utöver placeboeffekten).

Vi vill se en jämförelse med sedvanlig behandling, »usual care», för att veta om den nya behandlingen är lika bra eller bättre, och om biverkningar och pa-

tientpreferenser skiljer sig åt mellan alternativen. I praktiken visar det sig svårt att hitta sådana studier: vid osteoporos jämförs bifosfonater med kalk–D-vitamin, inte med östrogen; vid artros jämförs NSAID-preparat med placebo i stället för med paracetamol; de nya antidiabetiska glitazonerna jämförs med sulfonylurea, där en jämförelse med metformin vore det adekvata.

Det finns många exempel på att tillverkarna är rädda för att göra de kliniskt viktiga jämförande studierna, kanske på grund av att de kan visa att de nya behandlingarna är väsentligen likvärdiga med de gamla. Den sedvanliga behandlingen som ett läkemedel jämförs med behöver dock inte vara farmakologisk; vid diabetesstudier kan kost, motion och viktnedgång vara komponenter i en relevant kontrollgrupp, liksom socioterapi och minnesträning skulle kunna vara bästa jämförelsen vid demensstudier.

»Systerpreparaten»

Slutligen har vi problemet när ytterligare ett preparat, ett »systerpreparat», introduceras inom en grupp läkemedel. Ett nytt migränläkemedel inom triptangruppen kan visas vara bättre än placebo, och likvärdigt med bästa tidigare sedvanliga behandling (ASA/paracetamol plus metoklopramid), men någon direkt jämförelse med andra preparat inom gruppen presenteras inte.

Vi behöver åtminstone en jämförelse med originalsubstansen (sumatriptan), men får ofta nöja oss med data från farmakokinetiska studier eller indirekta jämförelser, där preparaten dock sällan testats mot varandra inom ramen för samma studie. Att hävda att det ena preparatet är bättre, eller ens likvärdigt, är då mycket svårt.

Vad vill jag uppnå med behandlingen?

Ett av de viktigaste tillskotten till klinisk medicin som EBM och medicinska metodutvärderingar har givit är en ny syn på hur vi visar att en behandling verkligen är till nytta för patienten. Ingen har ifrågasatt att syftet med behandling är att minska dödlighet och sjukdomskomplikationer, att lindra smärta och ångest, men forskningen har sällan visat att man verkligen har uppnått dessa mål. Det har ansetts tillräckligt att peka på att interventionen minskar någon intermediär variabel som man har förutsatt vara kausalt relaterad till död, komplikationer och lidande.

Användandet av sådana surrogat-effektmått har präglat forskningen fram till de senaste åren. Exempel är mätningar av »peak expiratory flow» (PEF) och »forced expiratory volume» (FEV)

vid obstruktiv lungsjukdom, blodtrycks- och kolesterolnivåer vid kardiovaskulär prevention, bentätetsmätning vid osteoporos, självrapporterad rökfrihet vid rökavvänjning, flebografiskt påvisade asymtomatiska ventromboser vid trombosprofylax, HbA1c vid diabetesstudier, flödesmätningar vid claudicatio, ultraljud av carotis vid strokepreventionsstudier eller koronarangiografisk regress vid lipidsänkning.

Problemet är att surrogat-effektmått sällan med någon precision verkligen förutspår kliniskt viktiga förändringar för patienten under behandlingen.

Utgångspunkten är patientrelaterad nytta

Vår utgångspunkt är att visa en patientrelaterad nytta av behandlingen. Då ställer vi oss inte frågan vad som är lätt eller spännande att mäta, utan reflekterar över vad vi och patienten på lång sikt verkligen vill uppnå med behandlingen. Förbättrad överlevnad är onekligen det mest konkreta utfallet vi kan tänka oss, men det är påtagligt sällan vare sig studerat eller visat i relation till många behandlingar.

Först på 1990-talet kunde vi faktiskt påvisa att primärpreventiv farmakologisk kolesterolsänkning leder till viss minskning av dödlighet; vad gäller förbättrad diabeteskontroll har nyttan av denna regim ännu inte kunnat bekräftas.

När vi formulerar vår fråga bör vi alltså reflektera över dels vilka risker patienten står inför på kort och på lång sikt, dels vilka inskränkningar i funktion och hälsorelaterad livskvalitet sjukdomen för med sig. Har verkligen de kliniska studierna visat att behandlingen har genomslagskraft avseende sådana viktiga effektmått?

När vi behandlar patienter med obstruktiv lungsjukdom kan vi inte nöja oss med förbättrade PEF-värden. Vi måste också kräva att behandlingen gynnsamt påverkar t ex exacerbationer ledande till akutbesök eller behov av tilläggsmediciner, behov av inläggning på sjukhus, sjukfrånvaro eller frånvaro från skolan. Dyspné, hosta, slembildning eller pip i bröstet besvärar patienten; de kan graderas i sjukdomsspecifika frågeformulär och mätningar av livskvaliteten. Med hjälp av sådana variabler hoppas vi bättre kunna fånga upp för patienten viktiga aspekter av sjukdomen, och därmed ge svar på frågan om effekten av och nyttan med behandlingen.

På samma sätt kan vi resonera när det gäller andra sjukdomar, men det är tyvärr osäkert om vi kan hitta studier som har ställt dessa frågor. Saknas kunskap om behandlingens effekt på viktiga patientrelaterade effektmått bör vi, pre-

cis som vid bristfällig forskningsdesign, vara återhållsamma med alltför tvingande slutsatser och rekommendationer.

Alla interventioner har någon möjlig ogynnsam effekt, och man bör ha funderat över vilka potentiellt viktiga biverkningar som skall vara belysta i de studier som skall hjälpa oss att besvara frågan om interventionens effekt. Sideffekter som leder till terapiavbrott kan leda till att nyttan i större patientgrupper blir begränsad, eftersom få patienter tål preparatet (som är fallet med akarbos vid diabetes). Större behov av kontroller på mottagningen eller laboratoriet kan också leda till att behandlingen blir omständlig och mindre attraktiv för patienten, kanske också dyrare för patient och samhälle.

En nyttig gruppövning

Den beskrivna processen kan tyckas omständig, men den har framgångsrikt prövats vid många gruppmöten med kolleger. Vi har träffats under en halv dag, diskuterat basala begrepp inom EBM och låtit samtliga deltagare beskriva ett vardagligt kliniskt problem som han eller hon inte hittat svaret på. Med stöd av hela gruppen har sedan problemet successivt omformulerats till en fyrdelad fråga (Ruta 4). Under denna interaktiva process får vi rikliga tillfällen att beröra alla ovanstående aspekter på hur en klinisk studie bör vara utförd för att ge det svar vi och patienten behöver.

En sådan fortbildningsinsats kan vara begränsad till enbart formuleringsdelen, men den kan också vara den första av fyra gruppmöten där varje möte berör en av punkterna i EBM-processen (Ruta 3). Jag har goda erfarenheter av flera studiecirkel med erfarna allmänläkare och sjukhusspecialister där man arbetat med sina egna kliniska problem för att slutligen kunna formulera ett kliniskt användbart svar.

Vår vardagliga kliniska verksamhet bör vara den största drivkraften för fortbildning och professionell utveckling. En process som söker evidensbasen för de kliniska beslutspunkterna kan vara ett stort steg framåt för varje enskild läkare att själv öka sin professionalism.

Referenser

- Bjørndal A, Flottorp S, Klovning A. Medisinsk kunnskapshandtering. Oslo: Gyldendal, 2000.
- Greenhalgh T. How to read a paper. The basics of evidence based medicine. London: BMJ Publishing Group, 1997.
- Sackett DL, Straus SE, Richardson WS, Rosenberg W, Haynes RB. Evidence-based medicine. How to practice and teach EBM. Second ed. Edinburgh: Churchill Livingstone, 2000.

Summary

Routine clinical decisions as an incitement to evidence-based medicine

Mats Eliasson

Läkartidningen 2000; 97: 3860–5.

Everyday clinical work raises many questions regarding diagnosis, prognosis and treatment. Traditional sources such as text books do not give all the answers, especially not in situations when decisions are needed without further delay. By learning the fundamentals of evidence-based medicine we may stand a better chance of founding our decisions on the best available evidence in order to provide better patient care. A focused question need to be posed, followed by a cost-effective literature search. This clinical question comprises of four parts: Who is the patient? Which is the intervention? What is the comparison? What outcome is relevant to the patient and what are the adverse effects?

Correspondence: Mats Eliasson, Sunderby sjukhus, SE-971 80 Luleå, Sweden.
E-mail: Mats.Eliasson@nll.se

Särtryck

LÄKARTIDNINGEN

Det stillasittande liv som präglar alltför många i de rika nationerna har negativa följder för hälsan. Fetma har blivit en folksjukdom, benskörhet blir allt vanligare och hjärt - kärlsystemet "äldras" snabbare än nödvändigt. Motion på recept är ofta en bra medicin.

Det är bakgrunden till den serie på 24 artiklar som Läkartidningen samlat i särtrycket av serien "Människan i rörelse". Den har tagits fram i samarbete med Svensk idrottsmedicinsk förening med syftet att redovisa de praktiskt viktiga erfarenheter och kunskaper som är av värde i läkarens vardagsarbete med inaktiva patienter, med motionärer som fått besvär av träningen, med astmatiker och gravida som undrar om de kan fortsätta motionera eller tävla. Här får läkaren hjälp att besvara de vanliga frågor som patienter ställer om motion och träningskador, och serien tar också upp de långsiktiga följderna samt etiska kontroverser kring idrott och idrottsmedicin. Priset är 150 kronor.

Människan i rörelse



Beställer härmed.....ex
av "Människan i rörelse"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till LÄKARTIDNINGEN
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

**ATT SÖKA,
GRANSKA OCH
VÄRDERA INFORMATION**

ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Den systematiska översikten grundval i evidensbaserad medicin

Ett av de viktigaste tillskotten till den kliniska medicinen under senaste årtiondet

Evidensbaserad medicin handlar om att fatta beslut om patienters utredning och behandling på basen av bästa tillgängliga vetenskap, sammanställd på ett systematiskt sätt. En central metod för kunskaps hantering inom den evidensbaserade medicinen är därför den systematiska översikten, systematic review. Begreppet härrör från en avgörande artikel av Cynthia Mulrow från 1987 [1]. Hon studerade 50 traditionella översiktsartiklar som under åren 1985–86 publicerats i fyra tunga medicinska tidskrifter. Ingen av översikterna uppfyllde alla uppställda kriterier på en god vetenskaplig syntes av tillgänglig kunskap.

De närmaste åren växte ett empiriskt bekräftat underlag fram för hur en systematisk översikt bör genomföras och hur den kan (och bör) kritiskt granskas [2]. De tio kriterier som används återges i Faktaruta 1. En grundtanke är att den som skriver en systematisk översikt följer samma vetenskapliga grundprinciper som forskarna som genomförde de ursprungliga studierna. En god systematisk översikt är därför en vetenskaplig produkt i sig, science on science, och skall redovisas på ett sådant sätt att frågeställning, metod, genomförande och tolkning är så klara att andra forskare kan reproducera resultaten.

Ett flertal jämförelser mellan traditionella översikter (narrative reviews) och systematiska sådana visar stor risk för systematiska fel (bias) och både över- och underskattning av behandlingseffekter i de icke-systematiska översikterna [3]. Som en följd av detta har under nittiotalet flera av de stora medicinska tidskrifterna, som JAMA, BMJ och Lancet, engagerat sig i den vetenskapliga diskussionen kring evidensbaserad medicin och kritisk granskning av forskning och publicerat

en mängd systematiska översikter. New England Journal of Medicine har där emot förhållit sig avvaktande och även publicerat mycket kritiska artiklar och ledare avseende metaanalyser och systematiska översikter. Den kritiska hållningen har till en del varit berättigad och tjänat att bromsa en alltför stor tillit till metoderna i sig.

Har kvaliteten i översikter förbättrats något sedan 1986? Nyligen publicerades en uppföljning av Mulrows undersökning omfattande 158 artiklar, varav endast två uppfyllde de tio krav som redovisas i Faktaruta 1 [4]. Knappt en fjärdedel redovisade hur litteratursökning utförts, hur originalstudierna granskats eller hur evidensen sammanvägts. Trots allt var resultatet betydligt förbättrat jämfört med tio år tidigare, men det visar också att vi har lång väg att gå innan översiktsartiklar i allmänhet är utförda med tillräckligt pålitlig metodik. Få uppfyller kraven på att kallas systematiska och på att utgöra en tillförlitlig grund för rekommendationer om behandling.

Den systematiska litteraturöversikten är ett av de viktigaste tillskotten till klinisk medicin under det senaste årtiondet. Därför har Läkartidningen beslutat att återkommande i kommenterad form presentera viktiga och välgjorda översikter.

Vi kommer också att med korta referat redovisa nya och spännande systematiska översikter och ge läsaren impulser att själv söka sig vidare och läsa dessa. (Internetadresser till översikterna ges i Faktaruta 2.)

De översikter som presenteras kommer huvudsakligen från följande tre källor:

1. systematiska översikter från the Cochrane Collaboration
2. systematiska översikter publicerade i referentbedömda tidskrifter
3. medicinska metodutvärderingar

Den evidensbaserade medicinen rymmer många nya sätt att se på vår kliniska verksamhet och på hur man överför forskningsresultat till mötet med den enskilde patienten eller till policybeslut, vårdprogram eller prioriteringsunderlag. I serien om evidensbaserad

Sammanfattat

- Evidensbaserad medicin kräver tillgång till bästa vetenskapliga underlag, sammanställt på ett systematiskt sätt.
- Traditionella översiktsartiklar redovisar sällan hela kunskapsbasen, innehåller ofta systematiska fel (bias) och överskattar mestadels behandlingseffekter.
- En systematisk översikt berör en fokuserad klinisk fråga, använder en strikt metodik för att finna, granska och kvalitetsgradera relevanta studier vars resultat oftast sammanfattas kvantitativt i en metaanalys.
- En välgjord systematisk översikt ger det bästa underlaget för att fatta beslut om riktlinjer för behandling av patienter men måste i det enskilda fallet tolkas i ljuset av klinisk erfarenhet och patientens preferenser.
- Cochrane Collaboration publicerar elektroniskt (CD-ROM och Internet) ständigt uppdaterade systematiska översikter som är fritt tillgängliga för Läkärförbundets medlemmar.
- De stora allmänna medicinska tidskrifterna publicerar i ökande utsträckning välgjorda och viktiga systematiska översikter, ofta fritt tillgängliga via Internet.
- Medicinska utvärderingsorganisationer, såsom SBU, sammanställer breda systematiska kunskapsunderlag, ofta kompletterade med hälsoekonomiska analyser.
- Medicinska tidskrifter och deras läsare bör ställa mycket höga krav på översikter.

Författare

MATS ELIASSON

med dr, överläkare, Medicin, Sunderby sjukhus, Luleå, projektsamordnare, Statens beredning för medicinsk utvärdering (SBU).

E-post: Mats.Eliasson@nll.se

FAKTARUTA 1

Tio kriterier på god vetenskaplig metodik i en systematisk översikt

1. Inriktad på en fokuserad fråga
2. Metoden för litteratursökning väl beskriven
3. Explicita kriterier för inklusion och exklusion av studier
4. Ingående studiernas metodik kvalitetsbedömd
5. Reproducerbar granskning och bedömning av originalstudier
6. Framtida forskningsbehov utpekade
7. Motstridiga resultat (heterogenitet) utforskade och tolkade
8. Kvantitativ syntes utförd där så är lämpligt och möjligt (metaanalys)
9. Viktiga patientrelaterade effektmått studerade
10. Fyndens generaliserbarhet diskuterad

medicin kommer läsaren att möta diskussion om ämnen som:

- Varför är randomisering, blindning och bortfall så viktiga i kliniska studier?
- Hur kan jag lära mig kritiskt granska vetenskapliga rapporter?
- Att läsa, förstå och förhålla sig kritiskt till metaanalyser
- Varför är surrogat-effektmått farliga och vad är egentligen de viktiga effekterna av en behandling ur patientens perspektiv?
- Varför är det vetenskapligt bedrägligt att publicera samma data vid flera tillfällen?
- När kan man lita på att det finns en generell klasseffekt av en grupp läkemedel?

Systematiska översikter i the Cochrane Library

Medlemmar i Läkarförbundet har numera fri tillgång till de databaser som ingår i the Cochrane Library. Via förbundets webbplats (www.slf.se) loggar man in med sitt medlemsnummer och efternamn på »Doktors sida» och klickar på »Cochrane» i den meny som finns till vänster. På webbsidan som kommer upp finns dels en svensk manual i att använda the Cochrane Library, dels en länk till dess webbplats.

Här följer några exempel på nyligen publicerade systematiska sammanställningar i the Cochrane Library. S k debriefing av personal som varit engagerad i traumatiska akutvårdsinsatser har blivit vanlig i svensk sjukvård. En sammanställning av åtta studier där personer som nyligen utsatts för en traumatisk händelse deltagit i individuell debriefing vid ett tillfälle visade ingen minskad risk för att utveckla posttraumatiska stressstillstånd, depression eller

FAKTARUTA 2

Internetadresser till källorna för de kommande översikterna

The Cochrane Collaboration	http://www.cochrane.org/
Läkarförbundet	http://www.slf.se/
JAMA	http://jama.ama-assn.org/
N Engl J Med	http://www.nejm.org/
BMJ	http://www.bmj.com/
Lancet	http://www.thelancet.com/
Ann Intern Med	http://www.acponline.org/journals/annals/annaltoc.htm
INAHTA (Internationellt nätverk för HTA)	http://www.inahta.org/
CCOHTA	http://www.ccohta.ca/
MSAC	http://www.health.gov.au/haf/msac/

ångest [5]. Författarna avråder från obligatorisk individuell debriefing vid traumatisituationer men tillägger att man inte värderat nyttan av gruppverksamhet. Sju studier av antituberkulös terapi vid Crohns sjukdom redovisas i en metaanalys som inkluderade 355 patienter [6]. Tillägg av antituberkulösa medel till standardbehandling ökade inte andelen patienter som bibehöll remission men i de två små studier som först använde steroider för att nå remission kunde betydligt fler patienter undvika nya skov. Steroidbehandling berörs också i en översikt om Guillain-Barrés syndrom omfattande 378 patienter i sex studier [7]. Efter fyra veckor var det ingen skillnad i neurologisk funktion mellan dem som fått steroidbehandling och dem som fått placebo.

The Cochrane Library har god täckning inom vården av svårt sjuka patienter, speciellt inom neonatologin där en aktuell översikt jämför dopamin- och dobutaminbehandling av hypotensiva för tidigt födda barn [8]. Man fann inga skillnader i dödlighet eller svåra cerebrala komplikationer men dopamin var mer effektivt för att höja blodtrycket, vilket dock bör uppfattas som ett surrogat-effektmått. De forskare som studerat effekten av intravenösa immunglobuliner vid sepsis eller septisk chock identifierade 23 relevanta studier som rapporterade dödlighet [9]. Metaanalysen påvisade en 40-procentig relativ riskreduktion för död med polyklont immunglobulin medan antiendotoxiner eller anticytokiner inte påverkade dödligheten. Den absoluta mortalitetsreduktionen var 4,2 procentenheter (från 38,2 till 34 procent), dvs 24 patienter måste behandlas för att rädda ett liv.

I the Cochrane Library hittar man även analyser av vardagliga krämpor och tillstånd. Nageltrång med åtföljande paronyki behandlas ofta med någon form av avlastning eller evulsio men har en tendens att återkomma. Lokalbehandling av nagelmatrix med fenol efter

evulsio minskade dramatiskt risken för recidiv jämfört med enbart evulsio men ledde till en påtaglig ökning av postoperativa komplikationer [10]. Illamående under tidig graviditet reducerades bra av antihistaminer, men också pyridoxin (vitamin B₆) och akupressur hade effekt jämfört med placebo [11]. Ultraljudsbehandling vid akut fotledsdorsion påverkade inte läkning och funktion efter sju dagar, enligt en metaanalys av fyra studier av modest kvalitet [12].

Påtagligt många översikter berör ämnen i utkanterna av skolmedicinen med oftast nedslående resultat vad gäller möjligheten att besvara frågan om effektivitet. Granskarna har oftast svårt att hitta några randomiserade och kontrollerade studier att basera sina slutsatser på. I dessa fall är det viktigt att hålla isär »avsaknad av bevis» från »bevis för avsaknad av effekt». Självklart skall frågan om huruvida behandlingen fungerar hållas öppen tills studier med adekvat design och statistisk styrka har presenterats. Holländska forskare studerade hypotesen att intag av vitlök skyddar mot arterioskleros [13]. Trots intensivt sökande efter studier hos fabrikanter och företrädare för vitlökssterapi lyckades man bara hitta en randomiserad studie av 78 patienter med claudicatio. Efter tolv veckor skilde sig gångsträckan inte mellan dem som fått vitlök och dem som fått placebo. Knappt en tredjedel av dem som fått vitlök klagade över lukten! Försöket att klarlägga ifall musikerterapi fyller en funktion vid behandling av dementia patienter misslyckades likaså då inga relevanta studier har publicerats [14]. Nästan lika illa var utfallet vid sammanställningen av antioxidanter och mineraltillskott för att förebygga makuladegeneration [15]. Endast en studie identifierades, och den påvisade inte någon effekt av preparaten. Studier av akupunktur vid rökning utgör ett undantag i detta sammanhang. Sammanlagt 18 studier inkluderas i en översikt som med stor tyngd drar slut-

satsen att akupunktur inte ökar sannolikheten för rökfrihet efter ett år [16].

Systematiska översikter i medicinska tidskrifter

Ett bra exempel på en översikt från »vanliga» medicinska tidskrifter berör orala antikoagulantia (OAK) vid kranskärlssjukdom där det fortfarande råder många oklarheter om nytta och risker. Till detta har bidragit att tidiga studier var små, övervakning av behandlingen med hjälp av INR (international normalized ratio) är obekvämt och blödningsrisken stor. Tillkomst av nya trombocyt-aggregationshämmare ledde också in-tresset bort från OAK, men trots att de nya preparaten börjat användas är frekvensen återinsjuknande och mortalitet vid kranskärlssjukdom ännu hög. En ny-publicerad noggrann metaanalys utgående från McMaster University inkluderade 31 randomiserade studier från 1960 fram till 1999, med minst tre månaders uppföljning, i syfte att klarlägga om OAK, med eller utan tillägg av acetylsalicylsyra (ASA), reducerar frekvensen av recidiverande hjärtinfarkt, slaganfall eller dödlighet och samtidigt väga in förekomsten av allvarliga blödningar [17].

Slutsatsen är att balansen mellan nytta och risk (i form av blödningar) är relaterad till intensiteten i behandlingen med OAK. Vid en jämförelse mellan lågintensiv antikoagulantibehandling (INR < 1,5) kombinerad med ASA och behandling med enbart ASA förebygger kombinationsbehandlingen 7 »händelser» (död, infarkt eller slaganfall) per 1 000 patienter men till en kostnad av 5 större blödningar. Måttlig till intensiv OAK-behandling (INR 2–4,8), jämfört med ASA, förebygger 13 händelser till priset av 14 blödningar, medan tillägg av ASA till OAK, i denna situation, förhindrar 54 händelser med 16 blödningar som följd. Jämförelser mellan OAK och placebo är av mindre relevans då ASA i dag är standardbehandling, men man kan notera att intensiv OAK-behandling (INR 2,8–4,8) förhindrade 98 händelser vid jämförelse med placebo, med 39 blödningar som följd. Författarnas slutsats är att kombinationen av moderat till intensiv OAK-behandling med ASA ter sig lovande med en modest blödningsrisk. Studier pågår för att närmare belysa denna behandling.

Medicinska utvärderingsorganisationers arbete

Att utvärdera medicinska metoder innebär oftast att sammanställa ett stort antal systematiska översikter inom en gemensam ram för att belysa en större frågeställning (t ex hur bör astma och KOL behandlas?) och att sedan komplettera sammanställningen med hälsoekonomiska och etiska analyser. I Sve-

rige representeras detta av SBUs arbete, och internationellt betecknas arbetssättet som Health Technology Assessment, (HTA). SBUs systerorganisationer publicerar många nyttiga rapporter, många på engelska eller med engelsk sammanfattning. Den kanadensiska HTA-organisationen redovisade 1999 en kritisk granskning av insulin lispro (»direktverkande») och fann inga övertygande belegg för en generellt förbättrad metabol kontroll trots lägre blodsockerstegringar efter måltid än vid användning av vanligt snabbinsulin [18]. Från Australien rapporteras att transmyokardiell laserbehandling av kranskärlssjukdom inte kan anses vara etablerad behandling med tanke på en perioperativ mortalitet på 3–5 procent och osäkra långtidseffekter [19].

Översikter bör alltid utföras och redovisas på ett systematiskt sätt, och medicinska tidskrifter bör kräva detta. Den läsare som vill ha en kunskaps-sammanställning med minsta risk för systematiska fel och största möjlighet att ge en realistisk bild av behandlingsnyttan bör söka efter översikter som uppfyller kriterier på systematik. Läkartidningen vill med denna serie bidra till att göra detta enklare.

Referenser

- Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987; 106: 485-8.
- Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. *Evidence-Based Medicine Working Group. JAMA* 1994; 272: 1367-71.
- Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomised control trials and recommendations of clinical experts. *Treatment for myocardial infarction. JAMA* 1992; 268: 240-8.
- McAlister FA, Clark HD, van Malraven C, Straus SE, Lawson FME, Moher D, Mulrow CD. The medical review article revisited: has the science improved? *Ann Intern Med* 1999; 131: 947-51.
- Wessely S, Rose S, Bisson J. Brief psychological interventions (»debriefing») for immediate trauma related symptoms and the prevention of post traumatic stress disorder (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Borgaonkar M, MacIntosh D, Fardy J, Simms L. Anti-tuberculous therapy for maintaining remission of Crohn's disease (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Hughes RAC, van der Meché FGA. Corticosteroid treatment for Guillain-Barré syndrome (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Subhedar NV, Shaw NJ. Dopamine versus dobutamine for hypotensive preterm neonates (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Alejandria MM, Lansang MA, Dans LF, Mantaring JBV. Intravenous immuno-

globulin for treating sepsis and septic shock (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.

- Rounding C, Hulm S. Surgical treatments for ingrowing toenails (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Jewell D, Young G. Interventions for nausea and vomiting in early pregnancy (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Van der Windt DAWM, Van der Heijden GJM, Van den Berg SGM, Ter Riet G, De Winter AF, Bouter LM. Ultrasound therapy for acute ankle sprains (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Jepson RG, Kleijnen J, Leng GC. Garlic for peripheral arterial occlusive disease (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Koger SM, Brotons M. Music therapy for dementia symptoms (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Evans JR, Henshaw K. Antioxidant vitamin and mineral supplementation for preventing age-related macular degeneration. (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- White AR, Rampes H, Ernst E. Acupuncture for smoking cessation (Cochrane Review). In: *The Cochrane Library*, Issue 4. Oxford: Update Software, 1999.
- Anand S, Yusuf S. Oral anticoagulant therapy in patients with coronary artery disease: a meta-analysis. *JAMA* 1999; 282: 2058-2067.
- Shukla VK, Otten N. Insulin lispro: a critical evaluation. Ottawa: Canadian Coordinating Office for Health Technology Assessment (CCOHTA), 1999.
- Transmyocardial laser revascularisation. Canberra: Medicare Services Advisory Committee (MSAC), 1999.

Summary

**Systematic review the foundation of evidence based medicine
One of the most important contributions to clinical medicine of the past decade**

Mats Eliasson

Läkartidningen 2000; 97: 2726-8.

The traditional narrative review has been shown repeatedly to be biased, mostly towards exaggerated treatment size effects. In contrast, the systematic review follows a strict protocol regarding focused questions, explicit criteria for literature searches, inclusion and exclusion criteria, critical appraisal and a synthesis which is quantitative when appropriate. There is empirical evidence that bias is reduced and that the conclusions reached have greater validity for the construction of treatment guidelines. In this paper three sources of systematic reviews are identified: the Cochrane Library, reviews published in peer-reviewed journals, and assessments of health technology (HTA-reports). Editors are encouraged to publish systematic reviews after proper critical appraisal, and readers are advised to search for such reviews when in need of guidance on important clinical questions.

Correspondence: Mats Eliasson, Dept of Medicine, Sunderby sjukhus, SE-971 80 Luleå, Sweden.

E-mail: Mats.Eliasson@nll.se

En guide till kritisk läsning av kliniska artiklar

Rötterna till den evidensbaserade medicinen kan sökas huvudsakligen vid McMaster-universitetet i Kanada där David Sackett och hans medarbetare fick en stor läsarskara med sin klassiska lärobok »Clinical epidemiology: A basic science for clinical medicine», sedermera uppdaterad [1]. Termen klinisk epidemiologi introducerades dock redan 1982 av makarna Fletcher, sedermera legendariska redaktörer för *Annals of Internal Medicine*. De beskriver begreppet klinisk epidemiologi som sprunget ur de två moderdisciplinerna klinisk medicin och epidemiologi [2]. Den är »klinisk därför att den försöker besvara kliniskt relevanta frågor och ge ledning vid kliniskt beslutsfattande med bästa tänkbara evidens». Den är »epidemiologisk därför att många av metoderna för att besvara dessa frågor har utvecklats av epidemiologer och därför att beslut om vården av enskilda patienter ses i en kontext av den stora populationen ur vilken patienten är hämtad». En viktig skandinavisk föregångare är dansken Henrik R Wulff, som redan 1976 publicerade boken »Rational diagnosis and treatment», nyligen utkommen i sin tredje upplaga [3].

Oxford-obstetriker lade grund till Cochranesamarbetet

Samtidigt attackerade epidemiologen Archie Cochrane de brittiska läkarna för bristande vetenskaplig stringens avseende bevisad nytta av de behandlingar man använde. Obstetrikerna i Oxford fängade upp utmaningen och byggde under 1980-talet upp den första systematiska sammanställningen av randomiserade och kontrollerade studier av behandlingar under graviditet och förlösning. Databasen blev en stor framgång, och när de brittiska hälsovårdsmyndigheterna 1992 avsåg att

stärka den vetenskapliga basen inom sjukvården gick uppdraget till Oxford. Till minnet av Archie Cochrane inrättades the Cochrane Center som sedermera växt till det världsomspännande Cochranesamarbetet. David Sackett flyttade till Oxford och prövade att omsätta klinisk epidemiologi i klinisk vardag på medicinavdelningarna vid Radcliff Infirmary och termen evidensbaserad medicin började sin vandring runt världen, med starkt stöd av Centre for evidence-based medicine i Oxford. På deras webbplatser (se rutan Internet-adresser) hittar man utmärkta exempel och material för egen fortbildning.

Genombrott med JAMAs artikelserie Users' guides ...

Kanadensiska och amerikanska läkartidningar var tidigt ute under 1980-talet med artikelserier som »Hur man läser kliniska tidskrifter» och »Hur man följer med i den medicinska litteraturen». Det stora genombrottet kom dock när *Journal of the American Medical Association (JAMA)* 1993 startade artikelserien *Users' guides to the medical literature* [4, 5]. En författargrupp med ursprung i McMaster-universitetet, Oxford och Cochranesamarbetet har sedan dess publicerat 27 artiklar i serien. Ämnen som avhandlas är hur man som läsare kan kritiskt granska och tillgodogöra sig vetenskapliga publikationer om bl a terapi och prevention [6-8], diagnostiska test [9, 10], systematiska översikter [11], kliniska riktlinjer och vårdprogram [12, 13], differentialdiagnos [14] och rekommendationer om screening [15]. I april i år publicerades den senaste delen, som är en vägledning till hur olika elektroniska databaser kan fungera som stöd för en evidensbaserad klinisk verksamhet [16].

Denna mycket välskrivna serie utgår från en klinisk vardag och vänder sig till läkare utan särskild forskningserfarenhet. Artiklarna ger enkla och robusta verktyg för den egna läsningen. Min egen erfarenhet är att nivån är väl anpassad för utbildningar, individuellt eller i grupp, av svenska kliniker. Jag vill speciellt rekommendera de två som berör terapi och prevention [6, 7]. Här finner man ett klokt resonemang kring många

Sammanfattat

- Dagens evidensbaserade medicin har rötter i klinisk epidemiologi i Nordamerika och i uppbyggnaden av systematiska databaser för kontrollerade kliniska studier i Oxford.
- Grunden för en evidensbaserad klinisk verksamhet är att på bästa tillgängliga vetenskapliga underlag fatta beslut om patienters utredning och behandling och samtidigt väga in egna erfarenheter och patientens preferenser.
- Att lära sig hela processen fram till ett evidensbaserat kunskapsunderlag kräver handledning, träning och mycket tid.
- Den sekundärlitteraturen, som kritiskt kommenterar sammanställningar eller enskilda viktiga studier, blir snabbt allt mer tillgänglig i elektronisk och uppdaterad form.
- Att erbjuda kliniskt verkliga kolleger lättillgängliga, ständigt uppdaterade och tillförlitliga kunskapsunderlag kan vara bästa vägen att uppnå bättre klinisk praxis och hälsa.

av den evidensbaserade medicinens fundament, såsom behov av randomisering och blindning, absoluta i stället för relativa riskreduktioner, konfidensintervall i stället för P-värden och behovet av att använda effektmått som är viktiga för patienten och inte bara surrogatvariabler. Glädjande nog pågår just nu ett arbete för att sammanställa och uppdatera alla artiklarna till en bok. I väntan på denna hittar man en bibliografi över denna serie och en mängd andra inom samma område på webbplatsen som

Författare

MATS ELIASSON

med dr, överläkare, Medicin, Sunderby sjukhus, Luleå, projektsamordnare, Statens beredning för medicinsk utvärdering (SBU).

E-post: Mats.Eliasson@nll.se

tillhör den italienska organisationen för evidensbaserad medicin, GIMBE.

För den som vill utvecklas från att bli översköld av den medicinska informationsfloden till att själv söka och hantera kliniskt relevant kunskap vill jag i första hand rekommendera »How to read a paper. The basics of evidence based medicine» av den engelska allmänläkaren Trisha Greenhalgh [17]. Det finns svenska erfarenheter av att använda boken i utbildningar allt från AT- och ST-läkare, allmänläkarnas FQ-grupper till SBUs expertgrupper, och alla är lika förtjusta i denna kloka och jordnära bok som ändå för ett strikt vetenskapligt resonemang. Boken kan beställas direkt från BMJ via Internet. Slutligen publicerades nyligen en norsk lärobok med samma inriktning [18].

Alla skall inte bli specialister på kritisk granskning

För att uppnå en medicinsk praktik som leder till vård av hög medicinsk kvalitet bör denna baseras på bästa tillgängliga vetenskapliga underlag. Idealiskt kan detta uppnås genom att träna alla läkare i den evidensbaserade medicinens fundament: formulera fokuserade frågor, sök svaren kostnadseffektivt, granska källorna kritiskt, syntetisera en slutsats och applicera kunskapen på ett omdömesgillt sätt i mötet med den enskilda patienten. Att uppnå dessa färdigheter kräver dock intensiva studier; handledd träning och genomförande är tidskrävande. I en aktuell ledare kommenterar några av den evidensbaserade medicinens portalgestalter att den ursprungliga ambitionen att föra ut arbetssättet till stora grupper knappast har visat sig vara en framkomlig väg [19].

Brittiska allmänläkare avvisar tanken att själva lära sig hela processen men välkomnar och använder gärna evidensbaserade kunskapssammanställningar i form av kliniska riktlinjer och vårdprogram. Min erfarenhet är att svenska läkare har liknade inställning – gärna evidensbaserad medicin men vi kan inte göra allt grundarbete själva.

Sammanställningar av kritiskt granskad vetenskap i form av antingen systematiska översikter eller kommenterade och strukturerade referat av centrala kliniska studier kan bli det redskap som efterfrågas. Denna sekundärlitteratur växer snabbt och blir allt mer tillgänglig, oftast via Internet [16]. Där finner vi först och främst the Cochrane Library (se LT 22/00). Den lokala »SBU-enheten» i Oxford ger ut en egen elektronisk tidskrift, Bandolier, med många kritiskt granskade och kliniskt viktiga ämnen kommenterade på ett lättillgängligt vis. På papper finns de två besläktade tidskrifterna Evidence-Based Medicine och ACP Journal Club, som kom-

Internet-adresser

Mac Master University, Clinical Epidemiology	www.fhs.mcmaster.ca/ceb/
Centre for Evidence-Based Medicine, Oxford	ceb.m.jr2.ox.ac.uk/
Cochranesamarbetet	www.cochrane.org/
JAMA	jama.ama-assn.org/
GIMBE (bibliografi över evidensbaserad medicin)	www.gimbe.org/Home.htm
BMJ	www.bmj.com/
Bandolier	www.ebando.com/
Best Evidence	www.acponline.org/catalog/electronic/best_evidence.htm
Clinical Evidence	www.evidence.org/index-welcome.htm

menterar centrala studier inom allmänmedicin, internmedicin, kirurgi och obstetrik. Just nu finns 1 600 kommentarer samlade på CD:n Best Evidence. I bokform utkom nyligen »Clinical evidence» (recenserad i LT 49/99), som kan ses som ett evidensbaserat terapikompendium uppdaterat två gånger per år och till hösten tillgängligt på Internet.

För att inte evidensbaserad medicin skall övertolkas bör begränsningarna i vardagen klargöras. Möjligheten att erbjuda kliniker lättillgänglig, kritiskt granskad och kommenterad kunskap kan bli den evidensbaserade medicinens viktigaste bidrag till bättre klinisk praxis och därmed bättre vård och hälsa.

Referenser

1. Sackett DL, Haynes RB, Guyatt GH, Tugwell P. Clinical epidemiology: A basic science for clinical medicine. Second ed. Boston: Little, Brown and Company; 1991.
2. Fletcher R, Fletcher S, Wagner E. Clinical epidemiology. The essentials. Third ed. Baltimore: Williams & Wilkins; 1996.
3. Wulff H, Götzche PC. Rational diagnosis and treatment. Evidence-based clinical decision making. Third ed. Oxford: Blackwell Science; 2000.
4. Guyatt GH, Rennie D. Users' guides to the medical literature [editorial]. JAMA 1993;270(17):2096-7.
5. Oxman AD, Sackett DL, Guyatt GH. Users' guides to the medical literature. I. How to get started. The Evidence-Based Medicine Working Group. JAMA 1993;270(17):2093-5.
6. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. JAMA 1993;270(21):2598-601.
7. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence Based Medicine Working Group. JAMA 1994;271(1):59-63.
8. Guyatt GH, Sinclair J, Cook DJ, Glasziou P. Users' guides to the medical literature: XVI. How to use a treatment recommendation. Evidence-Based Medicine Working Group and the Cochrane Applicability Methods Working Group. JAMA 1999;281(19):1836-43.
9. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence Based Medicine Working Group. JAMA 1994;271(9):703-7.
10. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence Based Medicine Working Group. JAMA 1994;271(5):389-91.
11. Oxman AD, Cook DJ, Guyatt GH. Users' guides to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group [see comments]. JAMA 1994;272(17):1367-71.
12. Wilson MC, Hayward RS, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. B. What are the recommendations and will they help you in caring for your patients? The Evidence-Based Medicine Working Group. JAMA 1995;274(20):1630-2.
13. Hayward RS, Wilson MC, Tunis SR, Bass EB, Guyatt G. Users' guides to the medical literature. VIII. How to use clinical practice guidelines. A. Are the recommendations valid? The Evidence-Based Medicine Working Group. JAMA 1995;274(7):570-4.
14. Richardson WS, Wilson MC, Guyatt GH, Cook DJ, Nishikawa J. Users' guides to the medical literature: XV. How to use an article about disease probability for differential diagnosis. Evidence-Based Medicine Working Group. JAMA 1999;281(13):1214-9.
15. Barratt A, Irwig L, Glasziou P, Cumming RG, Raffle A, Hicks N, et al. Users' guides to the medical literature: XVII. How to use guidelines and recommendations about screening. Evidence-Based Medicine Working Group. JAMA 1999;281(21):2029-34.
16. Hunt D, Jaeschke R, McKibbin K. Users' guides to the medical literature. XXI. Using electronic health information resources in evidence-based practice. The Evidence Based Medicine Working Group. JAMA 2000;283(14):1875-79.
17. Greenhalgh T. How to read a paper. The basics of evidence based medicine. London: BMJ Publishing Group; 1977.
18. Bjørndal A, Flottorp S, Klovning A. Medisinsk kunnskapshåndtering. Oslo: Gyldendal; 2000.
19. Guyatt G, Meade M, Jaeschke R, Cook D, Haynes R. Practitioners of evidence based care. Not all clinicians need to appraise evidence from scratch but all need some skill. BMJ 2000;320:954-955.

Cochranesamarbetets betydelse för evidensbaserad medicin

Cochranesamarbetet är en internationell organisation, med syfte att hjälpa bl a läkare och andra behandlare att fatta väl-informerade beslut om interventioner i sjukvården.

När man söker information om ett relevant kliniskt problem, kan man känna sig lite i samma situation som en bergsbestigare. Berget av litteratur är så stort, att det kan vara lika utmattande, dyrt och långvarigt att nå upp på toppen av Mount Evidence inom ett givet område som att bestiga Mount Everest. Den enskilda kliniker kan självklart inte klara att plöja igenom litteraturen varje gång han ställs inför en viktig fråga, men den nödvändiga informationen kan erhållas om vi delar arbetet mellan oss. Detta försöker man göra i det internationella Cochranesamarbetet.

Det har påpekats i den danska diskussionen att kardiologerna i jämförelse med gastroenterologerna sällan tvivlar om något, vilket de kanske borde göra i högre grad. Införandet av streptokinas i behandlingen av hjärtinfarkt påminner t ex om införandet av antibiotikabehandling vid magsår framkallat av *Helicobacter pylori*. Längre efter att evidensen konstaterats föreliggande fortsatte man att i läroböcker och översiktsartiklar skriva att streptokinas bara skulle användas vid speciella indikationer [1]. Behandlingen rekommenderades därför inte rutinmässigt förrän i slutet av 1980-talet, även om kunskapen uppnåddes omkring tio år tidigare. Man kan illustrera den ackumulerade evidensen

Författare

PETER C GÖTZSCHE

ledare för Det Nordiske Cochrane Center, H:S Rigshospitalet, Köpenhamn.

Artikeln är en redaktionellt bearbetad och uppdaterad sammanslagning av två artiklar som tidigare är publicerade i Ugeskr Laeger 1997; 159: 6721-4 respektive 1999; 161: 932-4.

med att göra en kumulativ metaanalys, där man varje gång man har en ny undersökning »lägger ihop den» med dem som tidigare publicerats. Då ser man att den första undersökningen inom området, som publicerades 1959, hade en mycket positiv effekt, men eftersom denna hade utförts på bara 23 patienter var resultatet behäftat med stor osäkerhet. Det är mycket karakteristiskt att de små, tidiga undersökningarna är alltför positiva, men efter hand får vi en mer stabil uppskattning av effekten av streptokinas.

Detta exempel är välkänt, och kanske känner några också till det motsatta exemplet, där man i läroböcker och översiktsartiklar fortsatte med att rutinmässigt rekommendera lidokain vid hjärtinfarkt, även om de kumulativa undersökningsresultaten under hela perioden ligger på den skadliga sidan. Även om effekten inte är *signifikant* skadlig, måste man säga att det är en besynnerlig diskrepans mellan rekommendationerna och evidensen.

Cochranesamarbetet

Cochranesamarbetet startade 1993 i Oxford. Det är en internationell organisation, vars syfte är att hjälpa läkare och andra behandlare, patienter, politiker och övriga aktörer att fatta välunderbyggda beslut om interventioner i hälso- och sjukvården.

Det gör vi framför allt genom att samla in randomiserade och kontrollerade studier (som vi handsöker i litteraturen), utföra metaanalyser utifrån dessa studier och uppdatera analyserna när det kommer fram nya studier. Cochranesamarbetet är uppkallat efter en brittisk epidemiolog, Archie Cochrane, som för många år sedan kritiserade professionen för att inte varje specialitet sammanställde sådana uppdaterade sammanfattningar av alla tillgängliga randomiserade studier.

En av huvudprinciperna för Cochranesamarbetet är samarbete framför konkurrens, bl a för att undvika dubbel forskning.

En annan princip är att Cochranedatabasen skall vara lätt tillgänglig, vilket vi försöker åstadkomma genom att hålla priset nere på drygt 1 500 kr om året,

»REDAKTIONSRTA»

Är du intresserad av att delta i Cochranesamarbetet? Du kan kontakta det nordiska Cochranecentret på e-postadress: general@cochrane.dk.

Du kan också kontakta den svenska representanten i centrets styrgrupp: Mats Eliasson vid Statens beredning för medicinsk utvärdering: Mats.Eliasson@nll.se tel 0920-28 20 00

Om det finns tillräckligt intresse kommer ett möte om att arbeta med Cochrane att anordnas i Stockholm i höst. Kurser i att skriva protokoll och genomföra metaanalyser i Cochranes programvara kommer i så fall också att arrangeras inom landet. Kontakta Mats Eliasson.

trots att databasen utkommer fyra gånger årligen.

Vi skall också vara uppdaterade, och det vi gör skall vara relevant för patienterna. Det rör sig därför inte så mycket om att undersöka effekten av våra behandlingar på surrogatmått, t ex blodtrycket eller andra markörer. Det handlar om ifall man överlever, om man mår bra och inte blir handikappad, vilket är sjukvårdens syfte.

Cochraneöversikt

Det ställs stora krav på en systematisk Cochraneöversikt. För att belysa principerna går jag kort igenom vår metaanalys av förebyggande av svampinfektioner hos cancerpatienter [2]. Det skall först och främst finnas en specifik frågeställning och en mycket utförlig datasökning, som bör beskrivas i detalj. Litteratursökningen innehåller därför flera sökord för randomiserade försök kombinerat med en rad preparat för svampbehandling och med en rad söktermer för immundefekter. Sökstrategin skall helst också beskriva att man har sökt publicerade arbeten genom att kontakta läkemedelsindustrin och aktiva forskare på området och kanske att man har bläddrat igenom konferenssammanfattningar osv. Det är mycket

sällan man i traditionella översiktsartiklar kan se hur stora ansträngningar författarna har gjort för att samla in litteraturen, och det saknas också specifika kriterier för vilka artiklar som skall inkluderas och för hur resultaten (evidensen) skall granskas kritiskt och därefter syntetiseras.

Översikten inleds med en sammanfattning och är i övrigt uppbyggd på sedvanligt sätt för vetenskapliga artiklar, t ex med ett bakgrundsavsnitt som beskriver vad det är man vill göra, vilka studier man söker efter och vilka patienter, interventioner, effektmål, sökstrategi och statistiska metoder man kommer att använda. Men så tillkommer något som går utöver många metaanalyser från tidskrifterna, nämligen en noggrann beskrivning av de enskilda studierna. I en separat tabell, som kan vara ganska omfattande, försöker man att beskriva vilken kvalitet de enskilda studierna har. I en annan tabell beskrivs de studier som inte är medtagna samt orsaken till att de inte ansågs relevanta eller kvalificerade till översikten, vilket är nog så viktigt.

Under resultaten kan man i en grafisk bild se t ex antal dödsfall och antal invasiva svampinfektioner uppdelade på de enskilda preparaten. Sist i översikten kommer en konklusion: Vad betyder resultaten för praxis och för eventuell framtida forskning?

Vi fann inte någon överbevisande effekt på mortaliteten, och bland de pati-

enter som överlevde skulle man behöva behandla 73 patienter för att förebygga endast ett fall av invasiv svampinfektion. Det tyckte vi var mycket liten utdelning av sådan behandling. Vi inkluderade därför att dessa medel skulle förbehållas patienter med dokumenterad infektion och patienter som deltog i fortsatta randomiserade försök.

Sådan blev vår slutsats men detta illustrerar också att slutsatser i någon grad är subjektiva, medan resultaten gärna skulle vara mer robusta. Det finns inte internationell konsensus inom området, och några anser kanske att de ändå önskar använda förebyggande svampbehandling, trots den ringa vinsten.

Svampöversikten är också publicerad i British Medical Journal [3], och det är mycket positivt att BMJ, Lancet och andra tidskrifter har ingått ett avtal om dubbelpublicering med Cochrane-samarbetet. Det är självklart alltid trevligt att bli publicerad i en av dessa framträdande tidskrifter, men å andra sidan är det mer tillfredsställande att publicera och uppdatera en Cochraneöversikt, då en papperspublikation snabbt blir föråldrad. Vår metaanalys blev t ex publicerad i april 1997, och det har redan kommit minst en ny relevant undersökning, varför analysen nu uppdaterats.

Det nordiska Cochrane-centret

I Cochranesamarbetet ingår 15 Cochrane-centra, som bl a har till uppgift att hjälpa forskare på alla upptänk-

liga sätt, t ex genom att arrangera arbetsmöten där principerna för systematiska översiktsartiklar går igenom. Vi har också en politisk funktion, framför allt med att skaffa pengar till Cochranesamarbetet, som är en ideell organisation. I synnerhet har National Health Service i Storbritannien gjort en mycket stor insats. Cochrane-centren har en rad speciella uppgifter, t ex utvecklar det nordiska centret den statistiska programvara som används till att göra Cochraneöversikter och vi koordinerar hela mjukvareutvecklingen.

Det nordiska Cochrane-centret (Det Nordiske Cochrane Center) med säte i Köpenhamn har ansvar för aktiviteter i de nordiska och de baltiska länderna, Polen, staterna i det tidigare Sovjetunionen och Mongoliet. Våra viktigaste samarbetspartner är UK Cochrane Centre i Oxford, BMJ, Statens beredning för medicinsk utvärdering (SBU), Afdelingen for Medicinsk Videnskabsteori vid Panum Institutet, Statens Institut for Medicinsk Teknologivurdering och Ugeskrift for Læger. Det nordiska Cochrane-centret har varit representerat i Cochranesamarbetets styrgrupp sedan starten 1993, och 1998 övertog ledaren för centrets norska gren ordförandeskapet för hela organisationen.

Norden står starkt i Cochranesamarbetet. Tre av de 47 internationella Coch-

Ingångssida till The Cochrane Library (<http://www.cochrane.dk/cochrane/cdsr.htm>).

What's new
About The Cochrane Library
How to subscribe
Frequently asked questions
Abstracts of Cochrane Reviews
Technical support
The Cochrane Collaboration
Internet subscribers
Information for contributors
Update Software

THE COCHRANE LIBRARY

The best single source of reliable evidence about the effects of health care

Available now: *The Cochrane Library 2000, Issue 2*

The Cochrane Library is published by *Update Software Ltd*
Please e-mail info@update.co.uk if you have questions or comments.

ranegrupperna har dansk ledning, nämligen Cochrane Hepato-Biliary Group, Cochrane Colorectal Cancer Group och Cochrane Anaesthesia Group, den sistnämnda täcker ett av de sista områden där det saknades en Cochrane-grupp. Många av grupperna har knutet till sig nordiska forskare som utarbetar Cochraneöversikter, och det finns nordiska redaktörer i nio av grupperna.

Metaanalys är en ung vetenskap och behovet av metodologisk forskning är stort. Också på detta område är nordiska forskare starkt representerade, genom att de bidrar till alla de elva existerande Cochranemetodgrupperna, som har till syfte att förbättra kvaliteten i Cochraneöversikterna. Tre av metodgrupperna har nordisk ledning, och det nordiska Cochranecentret har koordinerat arbetet i metodgrupperna och redigerat Cochranehandboken, som är en vägledning i att utarbeta Cochraneöversikter.

Många viktiga problemställningar, t ex biverkningar, går inte att eller är svåra att belysa i randomiserade försök. Det nordiska Cochranecentret har av den anledningen bildat en metodarbetsgrupp, som bl a skall utarbeta riktlinjer för när och på vilket sätt data från kohortstudier och andra icke-randomiserade sammanställningar kan användas i systematiska genomgångar av forskningslitteraturen.

De generella nordiska hälsovetenskapliga tidskrifterna är genomsköta efter randomiserade kliniska försök ända från 1948. Man har funnit 2 023 studier, av vilka omkring 1 000 inte var tillgängliga på Medline. Även om man antar att kostnaderna för att genomföra en klinisk studie bara är 100 000 kr, vilket är i underkant, motsvarar värdet av dessa »försunna» kliniska försök ca 100 miljoner kr.

I Cochranesamarbetet ingår 47 tvärfackliga översiktsgrupper. Som exempel kan nämnas grupperna som håller på med bröstcancer och lever-gallvägssjukdomar. Nästan alla sjukvårdens interventioner täcks av sådana översiktsgruppers intresseområden. Det som ännu inte är täckt försöker vi att klara genom att hänvisa till den av de existerande grupperna som kommer närmast det intresse som forskaren har. Inte bara forskare utan även t ex patientrepresentanter skall finnas i grupperna. Man förpliktar sig att uppdatera sin översikt när man väl fullföljt den. Både protokoll och metaanalyser referentbedöms innan de publiceras i databasen. Därefter startar en form av löpande referentbedömning, genom att de som använder databasen uppmanas komma med kritiska kommentarer, som författarna måste ta ställning till. En demokratiskt vald styrgrupp har den

övergripande beslutsrätten i Cochranesamarbetet.

Cochranedatabasen

Både i Cochranecentren och i översiktsgrupperna görs ett stort arbete med att gå igenom litteraturen genom att använda varje enskilt blad i tidskrifterna. Genom detta finner vi många studier som inte är registrerade i Medline och som vi lägger in i vår egen databas [4]. Denna databas innehåller för närvarande över 250 000 referenser till kontrollerade studier, men antagligen finns upp emot en halv miljon, eftersom många försök antingen är opublicerade eller bara finns som konferenssammanfattningar. Cochranesamarbetet har betytt att antalet försök som kan identifieras genom Medline under publikationstyperna *randomized controlled trial* eller *controlled clinical trial* har ökat från ca 20 000 till över 180 000.

I Cochranedatabasen finns för närvarande över 700 protokoll och lika många färdiga systematiska översikter. Utvecklingen var länge exponentiell med en fördubblingstid på bara tio månader. Även om tendensen nu försvagats kommer vi under de närmaste åren att nå upp i flera tusen Cochraneöversikter.

Cochranebiblioteket [4] innehåller flera andra värdefulla databaser. Om man t ex är intresserad av osteoporos kan man se att det också finns kommenterade systematiska översikter från papperslitteraturen och referenser till medicinska metodvärderingar (från SBU och dess internationella systerorganisationer) som annars är svåra att hitta.

Intresset stiger

Cochranesamarbetet bygger på de intressen forskarna har, och man kan anmäla sig för att göra en översikt om ett givet område. Man kan också delta genom att teckna sig för att söka igenom några årgångar tidskrifter för hand eller hjälpa till med att skaffa fondmedel. Jag hoppas att fler kliniska forskare kommer att ansluta sig. Vi får en del kontakter och vi kan märka att intresset accelererar, men när forskarna upptäcker hur mycket arbete som ligger bakom översikterna faller några dessvärre ifrån. Det är emellertid hög tid att komma till skott, eftersom ämnena blir upptagna. Det skulle vara förargligt att upptäcka att en person man kanske anser är mindre kvalificerad än man själv redan har lagt beslag på ens favoritämne. Om man inte själv har tid att göra arbetet kan man sätta igång några yngre forskare under sin vägledning och den grupp som forskarna kommer att tillhöra. Vi kan också från det nordiska Cochranecentret vara behjälpliga. Det finns en kortfattad introduktion om Cochranesamarbetet på Internet [5].

Referenser

1. Antman EM, Lau J, Kupelnick B, Mosteller F, Chalmers TC. A comparison of results of meta-analyses of randomized control trials and recommendations of clinical experts: treatments for myocardial infarction. *JAMA* 1992; 268: 240-8.
2. Gøtzsche PC, Johansen HK. Antifungal prophylactic or empiric therapy vs placebo or no treatment in cancer patients with neutropenia. The Cochrane Library. Issue 4. Oxford: Update Software, 1997.
3. Gøtzsche PC, Johansen HK. Meta-analysis of prophylactic or empirical antifungal treatment versus placebo or no treatment in patients with cancer complicated by neutropenia. *BMJ* 1997; 314: 1238-44.
4. The Cochrane Library [database on disk and CDROM]. The Cochrane Collaboration. Oxford: Update Software: 1996-. Updated quarterly.
5. <http://www.cochrane.dk> (Det nordiska Cochranecentrets hemsida, innehåller bl a en svensk broschyr om Cochranesamarbetet och har förbindelser till Cochranesamarbetets servrar).

Laparoskopisk kirurgi – hur evidensbaserad är den?



Modern laparoskopisk galloperation. Kirurgen assisteras av en operationsrobot.

Laparoskopisk kirurgi har inneburit en drastisk förändring inom kirurgin under 1990-talet, och man kan fråga sig vad som drivit och driver utvecklingen. Det finns många förklaringar till – och intressenter i – den snabba utvecklingen. Avsikten med det symposium som den här artikeln är ett referat av är att se tillbaka på erfarenhet och resultat av den laparoskopiska kirurgin. Det är förvisso inte första gången det görs, men

Författare

GÖRAN EKELUND

docent, överläkare, kirurgiska kliniken, Universitetssjukhuset MAS, Malmö/moderator vid symposiet

GUNNAR EDLUND

medicine doktor, överläkare, kirurgiska kliniken, Östersunds sjukhus, Östersund/gallkirurgiavsnittet

SAM SMEDBERG

medicine doktor, överläkare, kirurgiska kliniken, Helsingborg sjukhus, Helsingborg/bräckkirurgiavsnittet

CLAES RUDBERG

docent, överläkare, kirurgiska kliniken, Centrallasarettet Västerås, Västerås/appendektomiavsnittet

FOLKE JOHNSON

docent, överläkare, kirurgiska kliniken, Universitetssjukhuset Lund, Lund/refluxkirurgiavsnittet.

avsikten är att se var vi står idag och vad vi har för vetenskapligt stöd – »evidence» – och hur man ska tolka det.

Detta är angeläget inte minst för att verksamheten på sistone har ifrågasatts från flera håll. Således har nyligen såväl Danmarks motsvarighet till SBU som Danska institutet för klinisk epidemiologi [1] ställt sig bakom en utredning där man tyckte sig se att kolecystektomi med minilaparotomi, det vill säga med betydligt kortare bukväggssnitt än vid vanlig öppen operation, är ett säkert och ekonomiskt fördelaktigt alternativ till laparoskopisk kolecystektomi, men där man ändå konstaterade att här finns för lite studier och rekommenderade en hel rad randomiserade undersökningar. Vid Svensk kirurgisk förenings möte 1999 uppkom en rätt påtaglig debatt på samma tema. Nyligen har också frågan ställts om det inte är »dags att en gång för alla konkludera att laparoskopisk appendektomi är ute»[2]. Är det så? Eller har vi evidens för motsatsen? Ska vi fortsätta med vidareutveckling på den inslagna linjen? Behöver vi fler studier? I så fall vilka? Detta var vad symposiet på Svenska Läkarsällskapets riksstämma 1999 försökte besvara, med koncentration på fyra områden, som återges nedan.

Laparoskopisk gallkirurgi

Laparoskopisk kolecystektomi har på kort tid etablerats som normalmetod vid behandling av symtomgivande gallsten. Kritik har höjts för att metoden infördes okontrollerat [3]. Förespråkare

Sammanfattat

- Laparoskopisk kolecystektomi är säker, kostnadseffektiv och har god patientacceptans.
- Laparoskopisk ljumskbråcksplastik har en relativt lång »learning curve», är säker men dyrare än konventionell operation. Patientacceptans är god och infektionsfrekvensen lägre. Metoden är särskilt av värde vid dubbelsidiga bråck.
- Laparoskopisk appendektomi är säker, med lägre infektionsfrekvens. En något längre operationstid balanseras av kortare konvalescens. Den inledande laparoskopiska explorationen bedöms värdefull.
- Laparoskopisk fundoplicatio för refluxsjukdom är ännu ej fullt utvärderad. Resultaten hittills är goda vad gäller säkerhet, patientacceptans och vårdtid.
- Flera ytterligare frågor inom laparoskopisk kirurgi förtjänar att studeras.

för tekniken saknar dock inte entusiasm [4]. Vilka bevis – evidens – finns för laparoskopisk kolecystektomi? Försök att besvara frågan görs utifrån litteraturstudier med utgångspunkt från kontrollerade randomiserade studier (RCT), metaanalyser, andra systematiska rapporter och översiktsartiklar.

Randomiserade studier jämförande laparoskopisk och öppen kolecystektomi: I fyra studier [5-8] jämförs laparoskopisk kolecystektomi med traditionell öppen teknik. Ofta är det tämligen små patientserier. Samtliga studier tycks visa att laparoskopisk teknik är att föredra på grund av minskad postoperativ smärta, kortare vårdtid och

snabbare återgång till normal aktivitet.

Randomiserade studier jämförande laparoskopi och minilaparotomi: I tre större studier [9-11] ingår fler patienter än i ovan nämnda studier. En studie [11] är ofta citerad och den enda som är blind när det gäller utvärderingen. Den visade inte någon väsentlig skillnad mellan metoderna vad avser vårdtid eller sjukskrivning. Övriga studier [9, 10, 12] visade fördelar för laparoskopisk galloperation när det gäller vårdtid och smärta. Ett observandum i den förstnämnda studien [11] är den höga konverteringsfrekvensen på 20 procent. Detta kan tala för att kirurgerna inte var förtrogna med den laparoskopiska tekniken. Sammantaget tycks här redovisade studier visa att laparoskopisk teknik har fördelar eller är lika bra som minilaparotomi.

Koledokusskador: Ett allmänt problem vid utvärderingen av en ny teknik är att resultatet avspeglar en inlärningsprocess, »learning curve». Detta har belysts i bland annat en publikation [13] med genomgång av 8 839 fall. Koledokusskaderisken var betydligt större initialt, 1,7 procent jämfört med 0,17 procent efter erfarenhet av 50 operationer. Inlärningskurvan var av »låg tröskel»-typ, vilket innebär att kirurgerna under utbildning lär sig tämligen snabbt under förutsättning av tidig adekvat handledning. Koledokusskadeeffekvensen efter laparoskopisk kolecystektomi har varit föremål för omfattande studier, som dock försvåras av att rapporteringen är ofullständig och av att olika klassifikationer förekommer. Det framgår inte heller alltid klart om skadorna uppstått före eller efter eventuell konvertering. Skadefrekvensen är starkt varierande och har i en metaanalys för laparoskopisk kolecystektomi visats vara 3,6–4,7 promille och för öppen galloperation 1,9–2,9 promille [14]. Frekvensen tycks för laparoskopisk galloperation vara sjunkande.

Har redovisade studier tillräcklig bevismått? Randomiserade kontrollerade studier (RCT) anses ha högst bevisvärde, men en stötesten är att skillnader vad avser lågfrekventa komplikationer, som koledokusskada, knappast kan påvisas med någon högre grad av säkerhet [15]. Andra problem kan vara publikationsbias, selektion och osäkerhet beträffande kirurgernas förtrogenhet med respektive metod. Kunskaperna behöver således kompletteras [16] före ett ställningstagande.

Cochranesamarbetet, konsensusrapporter och översikter: I Cochranebiblioteket [17] finns ännu inte så mycket att hämta för den som söker evidens för laparoskopisk kolecystektomi. Dock återfinns här, från University of York, utvärdering av en omfattande me-

taanalys av större studier (över 100 patienter), där vissa kvalitetskrav måste vara uppfyllda för inklusion [14]. Totalt omfattar analysen 126 artiklar med sammanlagt 91 720 patienter.

Laparoskopisk kolecystektomi var föremål för 98 studier (78 747 patienter) och öppen kolecystektomi avhandlades i 28 studier (12 973 patienter). Det var stora variationer mellan studierna, men laparoskopisk operation var förenad med lägre mortalitet och något högre frekvens koledokusskador, dock sjunkande. Laparoskopisk kolecystektomi bedömdes i denna sammanställning som minst lika säker som öppen kirurgi, men trots det omfattande materialet konkluderar man att det finns utrymme för bättre upplagda studier, till exempel för att belysa kostnadsrelationer.

Som komplement till ovanstående kan nämnas konsensuskonferensen från National institutes of health (NIH) 1993 [18], en omfattande »review» från Royal college of surgeons 1996 [15], MFRs rapport 1997 [19] och en nyligen utkommen dansk rapport [1]. Det är inte helt lätt att sammanfatta dessa rapporter. Det finns uppenbarligen varierande uppfattning i vissa frågor. Några av de fördelar för laparoskopisk kolecystektomi som nämns är lägre mortalitet (evidens svagt), bättre postoperativ lungfunktion, mindre postoperativ smärta, snabbare återgång till aktivitet och kortare vårdtid.

I Sverige synes införandet av laparoskopisk kolecystektomi ha skett med god säkerhet och goda resultat [19]. MFR sammanfattar i sin rapport [19] att »retrospektivt förefaller det som om övergången till laparoskopisk operationsteknik varit lyckosam», men säger också »metodens kostnadseffektivitet i relation till minilaparotomi behöver ut-

värderas ytterligare», vilket även poängteras i de danska och engelska rapporterna [1, 15].

Konklusioner: Baserat på ovanstående genomgång bör följande slutsatser kunna dras:

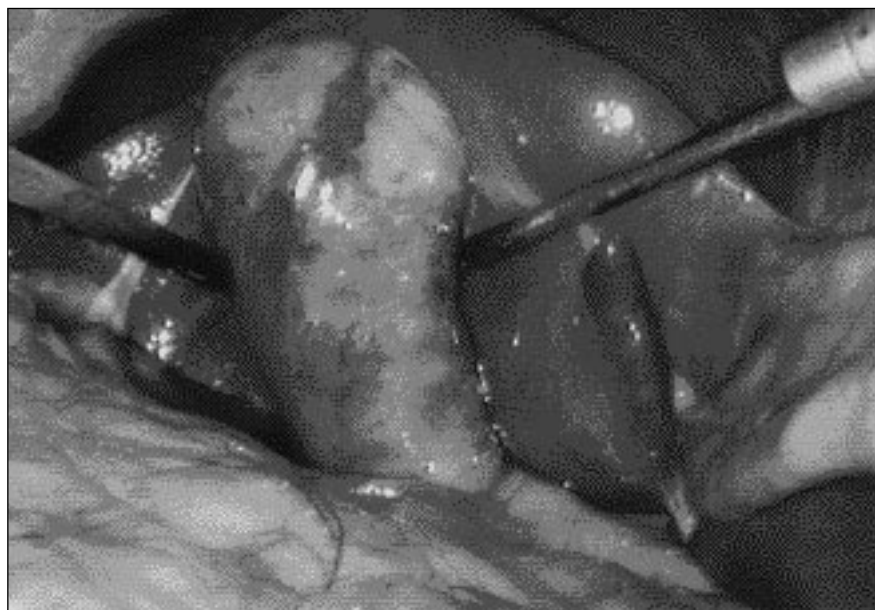
- laparoskopisk kolecystektomi är »evidensbaserad»,
- det finns ännu obesvarade frågor, till exempel gällande koledokusskadeeffekvens, teknik vid koledokuskirurgi och ekonomi,
- minikolecystektomi kan ge likvärdiga resultat, men dess »effectiveness» är mindre väl dokumenterad jämfört med laparoskopisk kolecystektomi.

Laparoskopisk bräckkirurgi

Bakgrund: De två senaste decennierna har bräckkirurgin genomgått en anmärkningsvärd utveckling. Nätanvändning, som är en nödvändig förutsättning för ett bra resultat vid laparoskopisk kirurgi, har även ökat kraftigt vid öppen bräckkirurgi. En jämförelse av teknikerna är inte endast en jämförelse mellan ny och gammal teknik utan återspeglar också frontlinjerna för olika utvecklingslinjer inom bräckkirurgin.

Operationsmetoder: Vid bräckoperation med konventionell plastisk friprepareras starka friska vävnadskanter utanför den av bräcket destruerade bakväggen i inguinalkanalen och sutureras. Av konventionella plastiker ger Shouldiceplastiken [20] bäst resultat. Resultatet av konventionell plastik vid recidivbräckoperation är klart sämre än vid operation för primära bräck.

Vid öppen nätplastik – främre och bakre – kan ett nät användas för att överbygga och förstärka området med defekt vävnad och därmed reducera eller helt eliminera tensionen i plastiken.



Laparoskopisk bild av kolecystit, med införda laparoskopiska instrument, inför kolecystektomi. Bild från Agneta Montgomery.

Den vanligaste främre metoden är Lichtensteinplastiken [21], som är lätt att lära och som ger bra resultat. Nätplastik förefaller ge lika bra resultat vid operation för recidivbräck som för primära bräck.

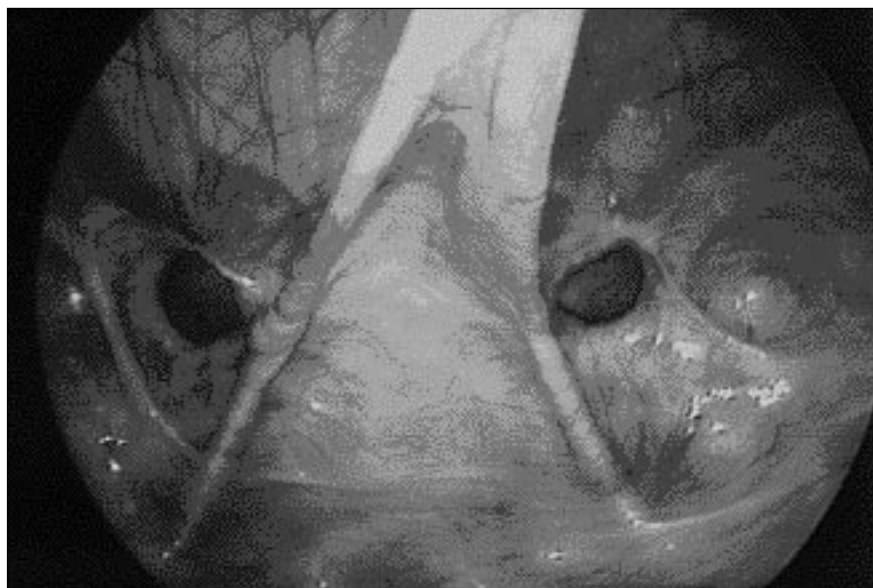
Med laparoskopisk teknik görs i princip samma operation som vid den öppna bakre nätplastiken. Nätet placeras på bukväggens insida, antingen med en transabdominell preperitoneal plastik (TAPP) [22] eller en total extra-peritoneal plastik (TEP) [23], varvid fossa iliaca och ljumskan nås med laparoskopet via rektusskidan som öppnas nedom naveln.

Resultat: De första randomiserade studierna av laparoskopisk bräckkirurgi publicerades 1994. Erfarenheten är således begränsad. De flesta studierna har kort uppföljningstid och de är fokuserade på peroperativa och tidiga postoperativa resultat som operationstid, komplikationer, postoperativa smärtor, sjukskrivningstider och kostnader. Recidivfrekvensen förväntas vara låg eftersom tekniken förutsätter nätanvändning, men långtidsresultat som bekräftar detta saknas fortfarande.

Denna genomgång omfattar 29 prospektiva randomiserade studier, publicerade före december 1999, jämförande laparoskopisk bräckoperation, TAPP respektive TEP, med öppna konventionella eller nätplastiker [24-52]. Materialet inkluderar totalt 2 677 laparoskopiska operationer och 2 857 öppna operationer.

Recidiv: Det viktigaste resultatet vid bräckkirurgi är recidivfrekvens. Studier inkluderande minst 200 randomiserade patienter med en uppföljning på minst ett år presenteras i Tabell I och II. En jämförelse mellan laparoskopisk och öppen konventionell bräckoperation (Tabell I) visar att recidivfrekvensen är densamma efter ett år, medan recidiven i den öppna gruppen blir fler redan under det andra året. Vid jämförelse med öppen nätoperation föreligger inga skillnader i recidiv under de första åren (Tabell II).

Komplikationer: Det »pris» patienten får betala för chansen till ett bra resultat med de olika metoderna är risk för komplikationer, smärtor och tid för tillfrisknande. Per- och postoperativa komplikationer har analyserats i 25 av studierna, inkluderande sammanlagt 5 534 patienter. Det fanns ingen mortalitet och komplikationerna är generellt mycket få. Den transabdominella laparoskopiska metoden TAPP har tillfört bräckkirurgin nya potentiella komplikationsrisker, som kärlskador och skador på inre organ. Frekvensen av dessa skador är dock så låg att den inte kan kalkyleras i randomiserade studier. Det man vet ifrån stora sammanställningar



Laparoskopisk bild av nedre delen av främre bukväggen utvisande bräckportar sedda från insidan. Bild från Agneta Montgomery.

är att en oerfaren kirurg och den första troakaren är de två största enskilda riskfaktorerna avseende skador i samband med laparoskopisk kirurgi [53]. Vid TEP är intraabdominella komplikationer i stort sett eliminerade, men då tekniken är tekniskt svår finns en ökad risk för konvertering till annan metod.

Sårinfektion förekom dubbelt så ofta efter öppen kirurgi (1,3 procent respektive 2,2 procent). Postoperativ neuralgi och kronisk smärta förekom också oftare efter öppen kirurgi (2,3 respektive 4,2 procent). Postoperativ konvalescens registrerades i 19 studier. Tiden för återgång till en normal aktivitetsnivå var kortare efter laparoskopisk operation i 16 studier, med en medeltid på 15 dagar jämfört med 21 dagar efter öppen kirurgi. Tre studier visade ingen skillnad.

Kostnaden för sjukvården: Operationstiden med laparoskopisk bräckoperation var i medeltal 68 minuter jämfört med 52 med öppen teknik. Vårdtiden redovisades i 24 studier. Dagkirurgi hade tillämpats i 10, och i övriga 14 studier var medelvårdtiden

2,5 dagar vid laparoskopisk kirurgi och 3,4 dagar vid öppen kirurgi. Kostnad för utrustning beräknades specifikt i 9 av studierna, och genomgående var den laparoskopiska operationen dyrare. I två svenska studier ökade kostnaden med 4 037 kronor [36] respektive 7 063 kr [34]. Användande av engångsmateriel fördubblar kostnaden vid en laparoskopisk bräckoperation.

Kostnad för samhället: Beräkningar av samhällets totalkostnad för bräckkirurgi i form av sjukvårdskostnad, sjukskrivningskostnader och produktionsbortfall är mycket svåra. I sex studier har beräkningar gjorts, tre av dem har samma författare, och man har kommit till olika resultat, till stor del på grund av variationer i sjukskrivningstider efter de olika operationsmetoderna. I en svensk studie resulterade den kortare sjukskrivningstiden efter laparoskopisk operation i en nettovinst på 11 392 kr [19], medan i en annan studie laparoskopipi istället medförde en nettoförlust på 5 988 kr eftersom skillnaderna i sjukskrivningstid här var mindre [34].

Tabell I. Recidiv. Antal efter ett respektive två år. Laparoskopisk operation kontra konventionell plastik.

Författare	Operationsmetod	Antal patienter	1 år	2 år
Dirksen (29)	TAPP	114		7
	Bassini	103		22
Johansson, SCUR (34)	TAPP	200	4	
	Kirurgens val	200	4	
Juul (35)	TAPP	138	4	
	Shouldice	130	3	
Kald (36)	Tapp	122		
	Shouldice	89	3	
Liem, COALA (41)	TEP	487	14	17
	Kirurgens val	507	15	31

Tabell II. Recidivfrekvens efter ett respektive tre år. Laparoskopisk operation kontra öppen nätteknik.

Författare	Operationsmetod	Antal patienter	1 år	3 år
Johansson, SCUR (34)	TAPP	200	4	
	Öppen preperit	200	11	
Khoury (37)	TEP	169		4
	Mesh-plug	146		4
The MRC trial group (49)	TEP (2/3)	468	7	
	Lichtenstein	460	0	

Konklusioner: Laparoskopisk bräckkirurgi kan, liksom övriga nätplastiker, förväntas ge lägre recidivfrekvenser än konventionella bräckplastiker på lång sikt. Komplikationsriskerna är likvärdiga. Patienttillfredsställelsen efter laparoskopisk bräckoperation är hög, med reducerade smärtor och snabbare tillfrisknande, men till priset av att sjukvårdens kostnader avsevärt ökas. Den relativa vinsten för samhället på grund av snabbare återgång till förvärvsarbete reduceras när sjukskrivningstiderna blir kortare även efter öppen kirurgi.

De första årens entusiasm och snabba ökning av antalet laparoskopiskt utförda bräckoperationer har planat ut. Fortfarande saknas långtidsresultat på vilka man mer säkert kan grunda rekommendationer för teknikens användning, men i väntan på dessa har hittills vunna erfarenheter givit en praxis. Laparoskopisk bräckkirurgi bör vara förbehållet de kirurger som har ett särskilt kunnande och intresse för laparoskopisk kirurgi. Vid samtidig bilateral bräckkirurgi blir det operativa traumat jämfört med ensidig operation obetydligt större med laparoskopisk teknik och innebär en vinst framför allt för patienten, men troligen också för sjukvården.

Laparoskopisk appendektomi

Bakgrund: Öppen appendektomi är en säker operationsmetod med mycket ringa mortalitet. Ett problem är fortfarande den relativt osäkra preoperativa diagnostiken liksom förekomsten av postoperativa komplikationer, där infektioner uppträder i 10–20 procent av fallen. Hos storvuxna patienter kan buksnittet behöva bli ganska stort, särskilt om appendix är svårtillgänglig.

Den laparoskopiska appendektomin är en minimalinvasiv operation som ger mycket god överblick i buken, vilket är av stort värde ur differentialdiagnostisk synpunkt. Frågan är då om den laparoskopiska tekniken kan ha fördelar beträffande postoperativ smärta, vårdtid, konvalescens och komplikationer liksom vad gäller total kostnad. För att skaffa vetenskapligt hållbara kunskaper om de olika metodernas effektivitet krävs prospektiva randomiserade studier,

korrekt utförda – bland annat enligt »intention-to-treat» principen och med en opartisk utvärdering av behandlingresultaten. I många studier syndas det mot basala krav, som en initial beräkning av materialstorleken baserad på förväntade skillnader mellan grupperna liksom inklusions- och exklusionskriterier med detaljbeskrivningar av hur operationerna utförts.

Resultat: Under de senaste 6–7 åren har en mängd studier gjorts för att belysa skillnader mellan öppen och laparoskopisk appendektomi, och de kvalitetsmässigt bästa studierna har inkluderats i flera metaanalyser. Problem här är att samma studie kan förekomma i flera av metaanalyserna. En sådan analys [54] omfattade 1 175 patienter i 10 studier och visade att laparoskopisk appendektomi hade längre operationstid, men något kortare vårdtid, kortare konvalescens och färre postoperativa infektioner. En annan sådan analys, från 1998, [55] innehöll resultat från 2 877 patienter ingående i 28 studier och visade att laparoskopisk appendektomi hade 16 minuter längre operationstid, något mindre postoperativ smärta, 15 timmar kortare vårdtid, 7 dagar kortare konvalescens och färre postoperativa sårinfektioner. År 1999 redovisades en metaanalys av 12 studier med 1 383 patienter [56]. Här visades 18 minuters längre operationstid för laparoskopisk metod, samma vårdtid, men 6 dagar kortare konvalescens och färre postoperativa infektioner. En metaanalys 1999 [57], omfattande 17 studier och 1 962 patienter, visade 15 minuters längre operationstid med laparoskopisk metod, men mindre postoperativ smärta, samma vårdtid, 6 dagars kortare konvalescens och färre postoperativa infektioner. Slutligen redovisade Fingerhut 1999 en metaanalys av 17 studier med cirka 1 800 patienter [2]. Här hade laparoskopisk appendektomi längre operationstid, något kortare vårdtid och kortare konvalescens. Konklusionen där var dock negativ mot den laparoskopiska tekniken, och kompletterades med en fråga om det inte var dags att konkludera att laparoskopisk appendektomi inte längre bör utföras. I Sverige har genomförts en multicenterstudie med 500

patienter, presenterad 1999, och det är för närvarande världens största studie [58]. Den visar att laparoskopisk appendektomi har 25 minuter längre operationstid, mindre postoperativ smärta och samma vårdtid, men 8 dagar kortare konvalescens.

Tolkning: När prospektiva randomiserade studier används inom kirurgin finns en del specifika problem. Om en ny operationsteknik ska jämföras med en gammal etablerad teknik finns det alltid en risk för att den nya tekniken inte görs rättvisa därför att man ännu inte fått tillräcklig träning med den nya metoden, det gäller alltså att gå in på rätt ställe på den så kallade inlärningskurvan. En annan aspekt är kirurgernas operationstekniska skicklighet. Opererar erfarna överläkare den ena gruppen och mindre erfarna underläkare den andra? Jämfördes den öppna och laparoskopiska metoden på lika villkor? Är den tekniska utrustningen tillräckligt bra? Och till sist, kan ny behandling ha så självklara fördelar att en randomiserad studie inte är nödvändig?

Konklusion: Konklusionen av denna genomgång av litteraturen är att evidens talar för att laparoskopisk appendektomi för närvarande har en längre operationstid, och därmed högre vårdkostnad, mindre postoperativ smärta, samma vårdtid, men kortare konvalescens, och därmed lägre samhällskostnad, och möjligen färre komplikationer. Om risken för ileus under de följande åren efter laparoskopisk appendektomi minskar får framtida studier utvisa.

Laparoskopisk antirefluxkirurgi.

Bakgrund: Laparoskopisk teknik vid antirefluxkirurgi har införts snabbt i Sverige. Redan 1996 utfördes två tredjedelar av antirefluxoperationer laparoskopiskt. Vid såväl öppen som laparoskopisk metod är principen att försnäva hiatus oesophagi genom att partiellt sy ihop crura diaphragmatica och utföra en så kallad fundoplikation. Syftet med denna genomgång är att analysera det vetenskapliga underlaget för övergång till laparoskopisk metod.

Resultat: Genom litteratursökning i Medline och Pubmed togs alla kliniska studier av laparoskopisk antirefluxkirurgi fram. I februari 2000 fanns 15 randomiserade studier, 21 kontrollerade, icke randomiserade studier och 65 okontrollerade studier. Av utrymmesskäl koncentreras framställningen till de 8 randomiserade studier som jämfört öppen och laparoskopisk kirurgi. De övriga 7 jämför olika laparoskopiska tekniker. Laine et al [59] fann, i en studie omfattande 110 patienter, stora skillnader i vårdtid och sjukskrivning till laparoskopins fördel. Vid uppfölj-

ning efter tre månader var det symptomatiska resultatet likvärdigt, även om man hade anmärkningsvärt hög frekvens patologiskt 24 timmars pH-värde i den öppna gruppen. Ytterligare en studie [60], från Finland och omfattande 42 patienter, visar stora skillnader i vårdtid och sjukskrivning. Detta gjorde att kostnaden för laparoskopin blev lägre. Även postoperativ smärta och trötthet var mindre i den laparoskopiska gruppen, men detta resulterade inte i någon skillnad i livskvalitet. I en studie från Lund på 60 patienter [61] gjordes en bländning av det postoperativa förloppet, vilket resulterade i att skillnaden i vårdtid blev betydligt mindre än i de båda finska studierna även om den fortfarande var signifikant kortare i den laparoskopiska gruppen. Dock fann man ingen skillnad i sjukskrivningstid. Vid uppföljning efter sex månader fann man inga signifikanta skillnader i det postoperativa resultatet, men frekvensen av mild dysfagi var högre i den laparoskopiska gruppen. Nyligen har en holländsk multicenterstudie [62] publicerats, där man avbröt inklusionen av patienter på grund av högre postoperativ dysfagifrekvens tre månader efter laparoskopi, 7/57 jämfört med 1/46. Två studier [63, 64] har jämfört immunförsvaret mellan patienter opererade med öppen respektive laparoskopisk metod, varvid man inte funnit någon skillnad frånsett mindre CRP-stegring efter laparoskopi.

Konklusion: Redovisad litteratur visar att laparoskopi ger kortare vårdtid och mindre postoperativ smärta, medan effekten på sjukskrivning och kostnad fortfarande är oklar. Det kirurgiska resultatet med båda metoderna förefaller på kort sikt likvärdigt, även om frågetecken måste sättas för dysfagifrekvensen. Det är utomordentligt viktigt att studier av långtidsresultat utförs.

Sammanfattande synpunkter

Trots att laparoskopin förts in mycket hastigt och utan föregående stora studier så tycks man i Sverige ha varit framsynt. Tidigt bildades ett laparoskopiregister och Svensk förening för laparoskopisk kirurgi har tagit på sig dels att testa nya produkter efter hand, dels att driva flera randomiserade studier. Laparoskopisk kolecystektomi kan mot en granskning i stora serier bedömas som säker och kostnadseffektiv, med god patientacceptans. Dock finns här fortsatt behov av utbildningsinsatser, standardisering och ackreditering. Ytterligare studier rörande patientacceptans behandlande koledokusten är önskvärda.

Vid laparoskopisk ljumskbråcksoperation synes postoperativ livskvalitet bättre vid laparoskopisk än vid öppen operation, men kostnaderna är på-

tagligt högre och den laparoskopiska bräckkirurgin är svår att lära, och det är tveksamt om vi idag har resurser för laparoskopisk operation, i varje fall av okomplicerade ensidiga ljumskbräck.

Laparoskopisk appendektomi kräver laparoskopiskt tränade kirurger, men tar ändå något längre tid än öppen kirurgi och är betydligt dyrare. Däremot finns det fördelar för patienter som hämtar sig snabbare, och infektionsfrekvensen är lägre. En stor fördel och vinst är den initiala explorativa laparoskopin.

Laparoskopisk operation för hiatusbräck kan ännu inte med säkerhet utvärderas eftersom antalet studier och antalet patienter här fortfarande är relativt litet. Hittills föreligger ingen större skillnad jämfört med öppen kirurgi, även om vårdtiden synes vara något kortare och den postoperativa smärtan mindre uttalad. Kostnadsrelationen är oklar.

Man kan således konstatera att laparoskopisk kirurgi har utvecklats ansvarsfullt i Sverige, men det finns mer att studera. Några steg tas nu tillbaka, men i stort tas många steg framåt.

Referenser

- Berggren U, Gordh T, Grama D, Haglund U, Rastad J, Arvidsson D. Laparoscopic versus open cholecystectomy: hospitalization, sick leave, analgesia and trauma responses. *Br J Surg* 1994; 81: 1362-5.
- Kiviluoto T, Sirén J, Luukkonen P, Kivilaakso E. Randomised trial of laparoscopic versus open cholecystectomy for acute and gangrenous cholecystitis. *Lancet* 1998; 351: 321-5.
- Majeed AW, Troy G, Nicholl JP, Smythe A, Reed MWR, Stoddard CJ et al. Randomized, prospective, single-blind comparison of laparoscopic versus small-incision cholecystectomy. *Lancet* 1996; 347: 989-94.
- Moore MJ, Bennett CL. The Southern Surgeons Club. The learning curve for laparoscopic cholecystectomy. *Am J Surg* 1995; 170: 55-9.
- Shea JA, Healey MJ, Berlin JA, Clarke JR, Malet PF, Staroscik RN et al. Mortality and complications associated with laparoscopic cholecystectomy. A meta-analysis. *Ann Surg* 1996; 224: 609-20.
- Downs SH, Black NA, Devlin HB, Royston CMS, Russell RCG. Systematic review of the effectiveness and safety of laparoscopic cholecystectomy. *Ann R Coll Surg Engl* 1996; 78: 241-323.
- Kavic MS. Laparoscopic hernia repair. Three-year experience. *Surg Endosc* 1995; 9: 12-5.
- McKernan JB, Laws HL. Laparoscopic repair of inguinal hernia using totally extraperitoneal prosthetic approach. *Surg Endosc* 1993; 7: 26-8.
- Johansson B, Hallerbäck B, Glise H, Anesten B, Smedberg S, Román J. Laparoscopic versus open preperitoneal mesh versus conventional technique for inguinal hernia repair. A randomized multicenter trial (SCUR hernia repair study). *Ann Surg* 1999; 230: 225-31.
- Kald A, Anderberg B, Carlsson P, Park PO, Smedh K. Surgical outcome and cost-minimisation-analysis of laparoscopic and open

hernia repair: A randomised prospective trial with one year follow-up. *Eur J Surg* 1997; 163: 505-10.

- Liem MS, van der Graaf Y, van Steensel CJ, Boelhouwer RU, Clevers GJ, Meijer WS et al. Comparison of conventional anterior surgery and laparoscopic surgery for inguinal-hernia repair. *N Engl J Med* 1997; 336: 1541-7.
- The MRC Laparoscopic Groin Hernia Trial Group. Laparoscopic versus open repair of groin hernia: a randomised comparison. *Lancet* 1999; 354: 185-90.
- McCall JL, Sharples K, Jadallah F. Systematic review of randomized controlled trials comparing laparoscopic with open appendectomy. *Br J Surg* 1997; 84: 1045-50.
- Sauerland S, Lefering R, Holthausen U, Neugebauer EA. Laparoscopic vs conventional appendectomy – a meta-analysis of randomised controlled trials. *Langenbecks Arch Surg* 1998; 383: 289-95.
- Temple LK, Litwin DE, McLeod RS. A meta-analysis of laparoscopic versus open appendectomy in patients suspected of having acute appendicitis. *Can J Surg* 1999; 42: 377-83.
- Chung R S, Rowland DY, Li P, Diaz J. A meta-analysis of randomized controlled trials of laparoscopic versus conventional appendectomy. *Am J Surg* 1999; 177: 250-6.
- Hellberg A, Rudberg C, Kullman E, Enochsson L, Fenyö G, Graffner H et al. Prospective randomized multicentre study of laparoscopic versus open appendectomy. *Br J Surg* 1999; 86: 48-53.
- Laine S, Rantala A, Gullichsen R, Ovaska J. Laparoscopic vs conventional Nissen fundoplication. *Surg Endosc* 1997; 11: 441-4.
- Heikkinen TJ, Haukipuro K, Koivukangas P, Sorasto A, Autio R, Sodervik H et al. Comparison of costs between laparoscopic and open Nissen fundoplication: a prospective randomized study with a 3-month followup. *J Am Coll Surg* 1999; 188: 368-76.
- Nilsson G, Larsson S, Johnsson F. Prospective randomized study between laparoscopic and open 360° fundoplication with blind evaluation. *Br J Surg*. Under publ.

En fullständig referenslista kan erhållas av Göran Ekelund, kirurgiska kliniken, Universitetssjukhuset MAS, 205 02 Malmö.

Summary

Laparoscopic surgery – evidence based at length; review and analysis of published results
Göran Ekelund, Gunnar Edlund, Sam Smedberg, Claes Rudberg, Folke Johnsson
Läkartidningen 2000; 97: 3457-62

The literature has been searched for current results in laparoscopic cholecystectomy, hernia repair, appendectomy and fundoplication. This was performed as a systematic review. Laparoscopic cholecystectomy was judged to be safe and cost/effective, with good patient acceptability. However a need for further studies is indicated. Laparoscopic technique in hernia repair has a longer learning curve and is more expensive than open repair, with no major difference in recurrence rates. It is preferable in bilateral repairs. Laparoscopic appendectomy in the hands of experienced surgeons is cost/effective. Time to recovery is shorter and the rate of infectious complications is lower than in conventional procedures. There are still too few results reported from laparoscopic fundoplication to permit reliable conclusions.

Correspondence: Göran Ekelund, Department of Surgery, Malmö University Hospital, SE-205 02 Malmö, Sweden.

E-mail: Goran.Ekelund@kir.mas.lu.se

Evidensbaserad komplementär medicin – finns det?

Första internationella kongressen diskuterar värde, säkerhet och risker

Väljer man att definiera alternativ, eller komplementär, medicin som »icke vetenskapligt prövade terapiformer» kan det förefalla finnas en inneboende motsättning i begreppet »evidensbaserad komplementär medicin». Ändå är det kanske föga förvånande att det ökande intresset för komplementär medicin nu kombinerats med den ökade betoningen på en evidensbaserad medicin inom den konventionella vården.

Vad är komplementär medicin?

Det finns ingen allmänt accepterad definition av vad komplementär medicin är. Det rör sig om terapiformer som vanligtvis inte lärs ut under läkarutbildningen, och som inte heller brukar erbjudas inom den slutna vården. Utvecklingen mot en ökad integrering av komplementär medicin inom hälso- och sjukvården har emellertid medfört att dessa redan vaga kriterier alltmer har förlorat i gränsdragande förmåga.

De terapiformer som diskuterades på kongressen, där över 200 läkare och forskare från fler än 20 länder – med tonvikt på USA, Storbritannien och Tyskland – samlats, var framför allt de i ett västerländskt perspektiv vanligast förekommande, och som också i stor utsträckning praktiseras av läkare: akupunktur, klassisk naturmedicin och homeopati, men även antroposofisk medicin och manuella behandlingsformer.

Utrymme gavs även åt behandlingsformer som ökat starkt i Europa och USA på senare tid, som Qi gong, shiat-su och healing.

Inte ett utforskat fält

Idag har Cochrane Collaboration Complementary Medicine Field registrerat ca 4 700 randomiserade kontrollerade studier (RCT) på området. Man har vidare samlat 204 systematiska översikter i en databas som omfattar 38 olika komplementära terapiformer. Det

Författare

MOTZI EKLÖF

doktorand, Tema hälsa och samhälle, Linköpings universitet.

E-post: motek@tema.liu.se

Sammanfattat

- De senaste studierna och metaanalyserna avseende området evidensbaserad komplementär medicin diskuterades den 6–8 april i München på den första internationella kongressen i ämnet.
- Ett särskilt tema var olika metodologiska svårigheter och utmaningar förknippade med forskning rörande komplementära terapiformer. Dessa problem gäller i stor utsträckning även forskning inom den konventionella medicinen.
- Ämnet är inte helt utforskat. Idag har Cochrane Collaboration Complementary Medicine Field registrerat ca 4 700 randomiserade kontrollerade studier på området.

handlar således inte om ett helt utforskat fält.

Professor Cynthia Mulrow, San Antonio Cochrane Center, definierade evidensbaserad medicin som »samvetsgrann, explicit och omdömesgill användning av det för närvarande bästa kunskapsunderlaget för beslutsfattandet inom hälso- och sjukvården». Hon ansåg att den komplementära medicinen mycket väl kunde vara evidensbaserad – om det är det man vill sträva efter.

Många av de frågor som diskuterades – som t ex värdet av randomiserade kontrollerade studier i förhållande till andra slags studier; hur den icke-specifika terapeutiska effekten skall kunna mätas på bästa sätt; hur den rätta balansen mellan vetenskap och läkekonst skall kunna uppnås – är gemensamma för den konventionella och den komplementära medicinen.

Utveckla omhändertagandet

Flera talare framhöll vikten av att utveckla omhändertagandet inom vården. Professor William Haskell, Stanford University, pekade på att det i USA idag satsas ungefär 100 miljoner dollar på forskning i komplementär medicin, men att motsvarande satsning inte görs för att utbilda läkarna till att bli bättre läkare.

David Reilly, Glasgows universitet och homeopatiska sjukhus, hörde till dem som betonade att det inte bara är

mer »evidence» som behövs inom medicinen utan även mer »confidence».

Nästa år igen

Kongressen hade organiserats av det komplementärmedicinska forskningscentret vid Münchens tekniska universitet, företrätt av Dieter Melchart, tillsammans med The Center for Alternative Medicine Research and Education vid Harvard Medical School, representerat av David Eisenberg.

Nästa kongress på temat kommer att hållas i San Francisco den 16–19 maj 2001.

Referenser

- Cochrane Collaboration Complementary Medicine Field Newsletter.
 Deutsche Zeitschrift für Akupunktur 2000; 43: nr 1 (temanummer samt abstracts från kongressen).
 Eklöf M. 1:a internationella kongressen om evidensbaserad komplementär medicin. ProSana: Tidskrift för komplementär medicin 2000; nr 2.
 Forschende Komplementärmedizin und Klassische Naturheilkunde 2000; 7: nr 1 (temanummer samt abstracts från kongressen).
 Röggla G. Complementary and alternative medicine at the crossroads? Lancet 2000; 355: 1343.

Duplikatpublicering ett sätt att försköna forskningsresultat

Oetiskt missbruk som hotar trovärdiga systematiska översikter och metaanalyser

Den systematiska översikten är ett bra instrument för att sammanställa samtliga tillgängliga studier rörande behandling av ett visst sjukdomstillstånd i syfte att ge en uppfattning om och i så fall vilka behandlingar som har effekt. Genom att numeriskt sammanväga storleken av effekten i de enskilda studierna kan en metaanalys presentera en kvantitativ uppskattning av den sannolika effektstorleken av behandlingen med betydligt större precision (smalare konfidensintervall) än enskilda studier förmår.

Vad är en metaanalys?

Jag har tidigare berört den systematiska översiktens kännetecken och vilka krav man som läsare bör ställa på en sådan översikt (Läkartidningen 22/00). En metaanalys skall ses som en utvidgning av den systematiska översikten. Metaanalysen är helt enkelt den matematiska sammanställningen av resultaten från den mödosamt uppbyggda systematiska översikten. Det innebär att man kan ha systematiska översikter utan metaanalyser, vilket oftast är fallet då de behandlingar som studerats skiljer sig mycket åt vad gäller preparat, doser eller behandlingstider, alternativt att sättet att mäta klinisk effekt är heterogent och svårt att på ett meningsfullt sätt väga samman. Men det innebär också att i den evidensbaserade världen kan en trovärdig metaanalys knappast presenteras utan att den baseras på en väl genomförd systematisk översikt.

Läkartidningen kommer under hösten och vintern att presentera flera artiklar med målsättningen att handleda läsarna i att läsa, tolka och förstå de stundtals svåröversiktliga presentationerna av metaanalyser. Vi vill upp-

Författare

MATS ELIASSON

med dr, överläkare, Sunderby sjukhus, Luleå; projektsamordnare, Statens beredning för medicinsk utvärdering (SBU); medicinsk redaktör (evidensbaserad medicin), Läkartidningen.

E-post: Mats.Eliasson@nll.se

Sammanfattat

- Studier som visar positiva effekter publiceras ibland i mer än en version, duplikatpublicering, ett vetenskapligt ohederligt och oetiskt förfarande, medan studier som inte visar positiva effekter publiceras mindre ofta.
- Om duplikatpublicering inte upptäcks kommer en metaanalys att beskriva en falskt förhöjd effekt av behandlingen och vilseleda klinikern.
- Prospektiv registrering av kliniska studier kan leda till att detta oskick stävjs.
- Exempel på duplikatpublicering avseende behandling med svampmedel, NSAID-preparat, risperidon, ondansetron och calcipotriol redovisas.

muntra till en kritisk hållning, och kommer därför att peka på de fallgropar som är vanligt förekommande i metaanalyser. Därmed undviker vi en förenklad syn på systematiska översikter och metaanalyser som varande ett slutligt svar på frågan om medicinska metoders effekt.

För att hindra den evidensbaserade medicinen från en, annars oundviklig, »backlash» är det viktigt att vi själva ständigt och öppet granskar våra egna instrument och utsagor.

Publikation av enbart positiva studier ger inflation av effekten

Varför är det så viktigt att fånga in så många studier som möjligt när vi systematiskt undersöker den kliniska effekten av behandlingar? Den viktigaste anledningen är att genom att täcka alla randomiserade och kontrollerade studier undviker man systematiska fel i tolkningen. Om enbart studier som uppvisat positiva resultat publiceras respektive identifieras som underlag för den systematiska sammanställningen leder det till stora brister i metaanalysen.

Denna »publication bias» leder till att den sammanvägda effektstorleken överskattas; man får en form av »inflation» i resultatets storlek och riktning. Om forskare, med eller utan stöd av anslagsgivande läkemedelsindustri, väljer

att stoppa studier med negativa resultat i byrålådan, och endast publicera positiva studier (och dessutom samma studie i många versioner), kan man snabbt visa att preparatet har god effekt.

Utbredd duplikatpublicering avseende flukonazolbehandling

Under arbetet med en systematisk litteraturoversikt avseende nyttan med antifungal medicinering vid neutropeni [1] stötte forskarna vid det nordiska Cochranecentret i Köpenhamn på oväntade problem [2]. Johansen och Gøtzsche fann 15 studier där dödligheten vid cancerassocierad leukopeni jämfördes mellan det nya preparatet flukonazol och amfotericin B. Knappt 80 procent av patienterna hade ingått i en jämförelse mellan flukonazol och peroral behandling av amfotericin B. Detta är mycket anmärkningsvärt, då peroral tillförsel, till skillnad från intravenös tillförsel, av amfotericin B inte har någon dokumenterad effekt vid detta tillstånd.

I en stor del av studierna var dessutom resultatet för amfotericin sammanslaget med resultatet för peroral nystatin (vilket inte heller har någon effekt på systemisk svampinfektion). Ett flertal rapporter omfattande knappt hälften av alla patienter liknade varandra mycket i studieupplägg, men smärre

detaljer i antal patienter och författarlistor skilde dem åt. I inget fall angavs att resultatet publicerats tidigare, men hela 92 procent av patienterna ingick i studier finansierade av tillverkaren av flukonazol, Pfizer Inc.

Detta väckte misstankar om att man dels publicerat samma multicenterstudie i ett flertal liknande skepnader, dels att man medvetet använt sig av en jämförelse som skulle visa att flukonazol var effektivare än amfotericin B. Parenthetiskt kan noteras att flukonazol inte var bättre i de få studier där man jämförde med intravenös tillförsel av amfotericin B.

Genom upprepade brev och andra påstötningar till såväl författarna av de enskilda studierna som tillverkaren försökte man klarlägga dessa misstankar, men utan resultat. Man noterade också att några forskare hade överlämnat sina originaldata till tillverkaren, och alltså inte längre hade tillgång till dessa. JAMAs redaktion erbjöd Pfizer att kommentera litteraturöversikten och de kritiska frågor som väckts, men företaget avböjde.

Vad visade Cochraneöversikten efter den komplicerade kartläggningen av vilka studier som egentligen var unika och skulle inkluderas? Hos patienter med neutropeni vid cancer var det bara amfotericin som minskade dödligheten, medan flukonazol och amfotericin i blygsam utsträckning minskade invasiv svampinfektion (men således utan effekt på dödlighet för flukonazol) [1].

Duplikatpublicering vanligt inom läkemedelsbehandling

I en kommenterande ledare ger Drummond Rennie, den medicinske redaktör som lett JAMA till en ledande ställning inom evidensbaserad medicin, tre andra exempel på sammanställningar av duplikatpubliceringar med förledande resultat [3]. En metaanalys av NSAID-preparat vid reumatoid artrit identifierade 31 studier som var publicerade sammanlagt 44 gånger. Sammanlagt 16 procent tolkades som duplikatpubliceringar, men i endast var fjärde av dessa angavs att resultaten tidigare var publicerade [4].

En aktuell sammanställning avseende det antipsykotiska preparatet risperidon utgick från 20 artiklar och en mängd opublicerade rapporter, men efter hårt arbete fanns två stora studier (varav en publicerad sex gånger) och sex små studier kvar att bygga metaanalysen på [5].

Effekten av ondansetron på postoperativt illamående beskrevs i 84 studier omfattande knappt 12 000 patienter, men efter granskning fanns bara 8 600 patienter i 70 studier [6]; 17 procent av studierna var duplikat, 28 procent av

patientdata rapporterades mer än en gång. Den allvarligaste följden av detta var en överskattning av behandlingseffekten med 23 procent.

En nyligen publicerad översikt om effekten av calcipotriol vid psoriasis identifierade 49 studier; tolv av dessa redovisade dock samma data som redan fanns i ingående studier, och de exkluderades därför [7]. I de fall man känner igen författarnamn, institutioner och studieupplägg kan man med rimlig arbetsinsats känna igen duplikaten, som ibland kan anses ganska oskyldiga – som de konferensabstrakt som redovisades i calcipotriolöversikten.

Många gånger är det dock svårt eller omöjligt ens för den som kan ämnet att se igenom de dimridåer som döljer det gemensamma ursprunget till många publikationer.

Avsteg från publicistisk norm

Det är uppenbart att publicering av sådana data är ett allvarligt avsteg från de principer för medicinsk publicering som uttryckligen förbjuder försök att få studier publicerade som redan har accepterats eller publicerats i annan tidskrift. JAMAs redaktör lägger ansvaret för detta missbruk dels på läkemedelsindustrin, som kortsiktigt hoppas kunna uppnå kommersiella fördelar, dels – och lika mycket – på de forskare som deltar i detta spel. Som positiva motkrafter utpekar han i första hand prospektiv registrering av påbörjade kliniska studier och publicering av resultat från alla studier, oavsett utfallet.

Register över påbörjade studier finns nu fritt tillgängliga på nätet: www.controlled-trials.com/

Att företaget som betalat studien sedan kan lägga hinder i vägen för att publicera negativa studier är väl belagt i ett stort antal fall, och uppfattas av många ledande kliniska forskare som ett tilltagande problem [8]. Arbetet i SBU:s projektgrupper har också klarlagt ett antal icke publicerade studier som i randomiserade undersökningar visat att olika preparat inte haft önskad effekt.

Sådan kunskap som erhållits via personlig kommunikation med dem som utfört ursprungsstudierna eller haft kännedom om »data on file» kan idag inte redovisas och användas i underlag för beslut om patienters vård och behandling.

Att inte publicera eller att dubbelpublicera är oetiskt

Kliniska studier som inte registrerats eller rapporterats i vetenskaplig litteratur kan inte bidra till att skapa den evidens som så väl behövs för att fatta beslut för hälso- och sjukvården. Det deklarerades vid en konferens som nyligen genomfördes av BMJ, Lancet och

brittiska läkemedelsföretag [9]. Självmant har Schering och Glaxo Wellcome beslutat att registrera och offentliggöra alla sina pågående studier inom ramen för Cochranesamarbetet.

Varför ytterligare läkemedelsbolag inte anslutit sig till detta initiativ har vi ingen förklaring till, men det torde vara den enda utvägen för de företag som vill betraktas som seriösa. Varför inte föreslå en automatisk registrering av alla svenska kliniska studier i samband med godkännande i etisk kommitté?

Vem kan man lita på?

Självfallet kan inte den enskilde läkaren kontrollera bakgrunden till de artiklar han/hon läser, och därmed bedöma om arbetet är ett duplikat. Inte heller kan man förstås veta något om de studier som inte publicerats. Det är här som de systematiska översikterna kan komma till hjälp. Genom att noggrant fånga upp publicerade studier och noggrant granska dessa är det ofta (men inte alltid) uppenbart vad som har publicerats mer än en gång, och därmed kan man begränsa sina analyser till de ursprungliga studierna.

Slutsatserna från systematiska sammanställningar riskerar att i mindre utsträckning drabbas av systematiska fel. Behandlande läkare och patienter kan därför vara tryggare i att beslut om behandling fattas på basen av för närvarande bästa tillgängliga kunskap, dvs evidensbaserad medicin.

Därför är en av hörnstenarna i Läkartidningens satsning på evidensbaserad medicin att publicera och kommentera systematiska översikter inom ämnen med relevans för svenska kliniker. Vi kommer att bevaka att sådana översikter har använt bästa tänkbara metoder och stringens för att läsarna skall kunna lita på att det är just bästa tillgängliga evidens vi erbjuder.

Översikter från Cochranesamarbetet och de medicinska metodvärderingsorganisationerna (dvs SBU-liknande organisationer), liksom många publicerade översikter i medicinska tidskrifter, uppfyller oftast dessa krav, och utgör därför basen i vår satsning på evidensbaserad medicin.

Referenser

1. Gøtzsche P, Johansen H. Routine versus selective antifungal administration for control of fungal infections in patients with cancer (Cochrane Review). In: The Cochrane Library, Issue 2, 2000. Oxford: Update Software.
2. Johansen HK, Gøtzsche PC. Problems in the design and reporting of antifungal agents encountered during meta-analysis. *JAMA* 1999;282:1752-9
<http://jama.ama-assn.org/issues/v282n18/full/jsc90228.html>
3. Rennie D. Fair conduct and fair reporting of

- clinical trials. JAMA 1999;282:1766-8. (<http://jama.amaassn.org/issues/v282n18/full/jed90083.html>).
4. Gøtzsche PC. Multiple publication of reports of drug trials. Eur J Clin Pharmacol 1989;36:429-32.
5. Huston P, Moher D. Redundancy, disaggregation and the integrity of medical research. Lancet 1996;347:1024-6.
6. Tramèr MR, Reynolds DJM, Moore RA, McQuay HJ. Impact of covert duplicate publication on meta-analysis: a case study. BMJ 1997; 315: 635-640 (<http://www.bmj.com/cgi/content/full/315/7109/635>).
7. Ashcroft D, Po A, Williams H, Griffiths C. Systematic review of comparative efficacy and tolerability of calipotriol in treating chronic plaque psoriasis. BMJ 2000; 320:963-7 (<http://www.bmj.com/cgi/content/full/320/7240/963>).
8. Bodenheimer T. Uneasy alliance. Clinical investigators and the pharmaceutical industry. N Engl J Med 2000;342:1539-44 (<http://www.nejm.org/content/2000/0342/0020/1539.asp>).
9. Tonks A. Registering clinical trials. BMJ 1999;319:1565-8 (<http://www.bmj.com/cgi/content/full/319/7224/1565>).

Summary

Duplicate publication is endangering the validity of systematic reviews and meta-analysis

Mats Eliasson

Läkartidningen 2000; 97: 3454-6

Studies showing favourable results tend to be published more than once. This practice of duplicate publication is, however, dishonest and unethical on the part of scientists and pharmaceutical companies. Inclusion into a meta-analysis of the results from multiple papers derived from one and the same controlled study gives an inflated picture of the efficacy of the treatment.

If manufacturers and researchers not only made all their study data public but also registered trials prospectively these pitfalls could be avoided.

Correspondence: Mats Eliasson, Sunderby sjukhus, SE-971 80 Luleå, Sweden.
E-mail: Mats.Eliasson@nll.se

SÄRTRYCK ur LÄKARTIDNINGEN

■ TILLVÄXTFAKTORER

Alla kroppens celler reagerar på olika signalämnen i omgivningen. De kallas tillväxtfaktorer. 12 artiklar speglar forskning och tillämpning. 56 sidor. 90 kr.

■ MISSÖDEN, MISSTAG, MISSBRUK

Hur löser man konflikter vid missöden i vården? 21 artiklar om problemläkare, läkarproblem, ansvarsfrågor och patientförsäkring. 80 sidor. 75 kr.

■ VÅLD OCH AGGRESSIVITET

Våldet möter läkare på skilda nivåer inom vården. Det kan också drabba dem själva. 26 artiklar ger ett brett perspektiv på våld och aggressivitet. 84 sidor. 95 kr.

■ ENLIGT MIN ERFARENHET

32 korta, praktiskt inriktade artiklar med anknytning till vårdens vardag. Diagnostik, terapi, exempel på prevention och ledningsfrågor. 48 sidor. 55 kr.

■ LUNGCANCER

Kunskaperna om lungcancersns biologi har ökat väsentligt på senare år, vilket innebär nya möjligheter för både prevention och behandling. Kombinationen av flera terapeutiska principer innebär bot, eller lindring, för fler patienter. Sex artiklar ger överblick över möjligheter och begränsningar med dagens terapimetoder. 36 sidor. 60 kr.

Härmed beställs

..... ex Tillväxtfaktorer

..... ex Missöden, misstag, missbruk

..... ex Våld och aggressivitet

..... ex Lungcancer

..... ex Enligt min erfarenhet

Namn

Adress

Insändes till Läkartidningen, Box 5603, 114 86 Stockholm. Telefax 08 - 20 76 19

Rapport från konferens

Evidensbaserad medicin med stöd av kvalitativ forskning

Den randomiserade kliniska prövningen har kommit att bli den evidensbaserade hälso- och sjukvårdens gyllene standard. Detta vetenskapliga paradigmen har genom sina strikta procedurer många starka sidor och förtjänster. Emellertid har det också svagheter. Dessa har framför allt att göra med att paradigmet utgår från tanken att människan bäst förstås och analyseras som ett objekt. Det bortser i stor utsträckning från att människor är subjekt med önskningsar, känslor, avsikter och handlingar.

Viktigt att utveckla den praktiska tillämpningen

Insikten om dessa begränsningar har lett till ett intresse för att använda så kvalitativa forskningsmetoder även inom läkarvetenskapen; först allmänmedicinen [1, 2] och senare också de sjukhusbundna specialiteterna [3]. Detta gör det viktigt att utveckla den praktiska tillämpningen av den kvalitativa forskningens resultat så att standarden för patienters behandling och vård vidmakthålls och förbättras. Erhållna kunskaper och insikter bör gå att värdera genom perspektivet evidensbaserad medicin, vilket Lars Werkö [4] träffande definierat som »en praktisk filosofi, som innebär att varje åtgärd i sjukvården skall vara baserad på bästa möjliga tillgängliga kunskap».

Premiär för konferens

Mot denna bakgrund anordnades vid Coventry Universitet under maj i år den första konferensen någonsin om »qualitative evidence-based practice». Den var välbesökt och givande med plenarföreläsningar, stimulerande workshops och presentationer av enskilda forskningsarbeten.

Tematiskt täckte den ett brett och spretigt område. Inom detta kunde man

urskilja väsentligen tre egenskaper, som karakteriserar de kvalitativa metoderna och skiljer dem från de kvantitativa. För det första fångar de upp och strukturerar händelser/fenomen utifrån de studerade personernas perspektiv, för det andra bidrar de med beskrivningar av de komplexa sammanhang vari dessa ingår, och slutligen betonar de helheter och processer.

Värdefulla kvalitativa metoder

Detta gör att de kan bidra till svar på:

1. Sammanhangsrelaterade frågor, att identifiera existerande former och väsen, t ex hur människor uttrycker erfarenheter och behov av hälso- och sjukvård eller hur de ordnar och tolkar symtom i relation till sammanhang (jämför t ex »modern age diseases» [5]).
2. Diagnostiska frågor, att undersöka skälen till varför man använder existerande former. Varför uppstår t ex behov av att utnyttja den ena eller andra vårdformen, varför står mängder av ordinerade läkemedel oanvända i människors skåp?
3. Utvärderande frågor, att uppskatta effektiviteten av existerande former, t ex vad avgör om en erbjuden form accepteras?
4. Strategiska frågor, att identifiera/utveckla nya teorier, planer eller ny vårdpolitik, t ex vilka typer/former av hälso- och sjukvårdstjänster kan tänkas motsvara uttryckta behov?

Kvalitativa forskningsmetoder ett komplement

Frågorna kan inte alltid besvaras enbart med kvalitativa metoder, men de är nödvändiga, självständiga komplement till andra procedurer.

Deras praktiska användbarhet ökar genom förbättrad kunskap om hur de kan värderas. För att bidra till en sådan utveckling skall konferensen återkomma årligen. Många professioner var representerade i år, dock bara två läkare, varför jag hoppas att betydligt fler deltar nästa gång i maj år 2001.

Information kan erhållas via Internet-adresserna m.c.taylor@coventry.ac.uk och m.savinbaden@coventry.ac.uk (tel +44-(0)24-768 87 04 50).

Referenser

1. Rudebeck CE. General practice and the dialogue of clinical practice. On symptoms, symptom presentations, and bodily empathy [dissertation]. *Scand J Prim Health Care* 1992; suppl 1.
2. Johansson E. Beyond frustration: Understanding women with undefined musculoskeletal pain who consult primary care [dissertation]. Umeå University medical dissertations; New Series No 550. Umeå: Umeå Universitet, 1998.
3. Bunne M. Eustachian tube function and tympanic membrane retraction. Qualitative and quantitative studies on variability and on related symptoms [dissertation]. Umeå University medical dissertations; New Series No 646. Umeå: Umeå Universitet, 1999.
4. Werkö L. Evidensbaserad medicin. Ett konkret redskap för vårdens verkstadsgolv. *Läkartidningen* 2000; 97: 2710-3.
5. Arnetz BB. Causes of change in the health of populations: A biopsychosocial viewpoint. *Soc Sci Med* 1996; 43: 605-8.

Författare

ÅKE THÖRN

med dr, överläkare, yrkesmedicinska enheten, Sunderby sjukhus, Luleå.

E-post: ake.thorn@nll.se

Det är lätt att söka medicinsk information på Internet

Så här gör du – steg för steg

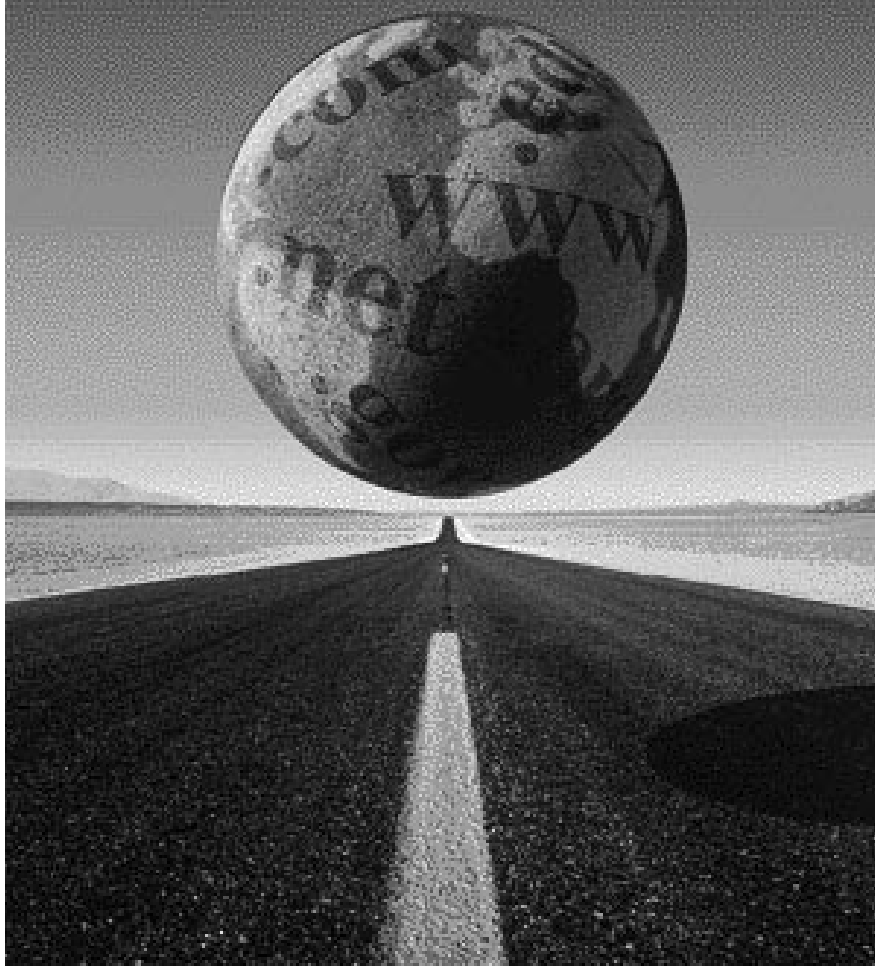


FOTO: CHARLY FRANKLIN/IBEL

Sammanfattat

- Kliniska beslut kräver tillförlitliga, lätt tillgängliga och ständigt uppdaterade kunskapskällor.
- Systematiska översikter finns i stor utsträckning på Internet.
- Clinical Evidence och Cochrane-biblioteket är bra startpunkter.
- Medline-sökning är ett komplement att använda först när bra översikter saknas.

Författare

VIVEKA ALTON

informatiker, Statens beredning för medicinsk utvärdering (SBU), samordnare för SBUs osteoporosprojekt

E-post: alton@sbu.se

MATS ELIASSON

med dr, överläkare, Sunderby sjukhus och SBU; medicinsk redaktör (evidensbaserad medicin) Läkartidningen.

Evidensbaserad medicin innebär att varje åtgärd i sjukvården skall vara baserad på bästa möjliga tillgängliga kunskap. Det gäller för den enskilde läkaren att hålla kunskapsförrådet aktuellt, vilket inte alltid är så lätt för den som står mitt uppe i rutinsjukvårdens grottekvart [1].

Vi vill i denna artikel visa hur man från en klinisk problemställning kan söka sig fram i olika informationskällor, tillgängliga via Internet, till en användbar kunskapssammanställning.

Före sökningen: Fokusera frågeställningen!

Det första vi måste göra är att tydliggöra och analysera frågeställningen, så att vi verkligen vet vad vi vill ha svar på. Vi har i en tidigare artikel berört frågan om att gå från det mer diffusa kliniska problemet till en fokuserad frågeställning [2]. För denna genomgång har vi valt problemställningen »behandling av osteoporos med bifosfonater», och för att besvara frågan om en behandlings effekter är det naturligt att vi söker efter systematiska litteraturöversikter, meta-analyser eller randomiserade, kontrollerade studier, »randomised controlled trials» (RCT).

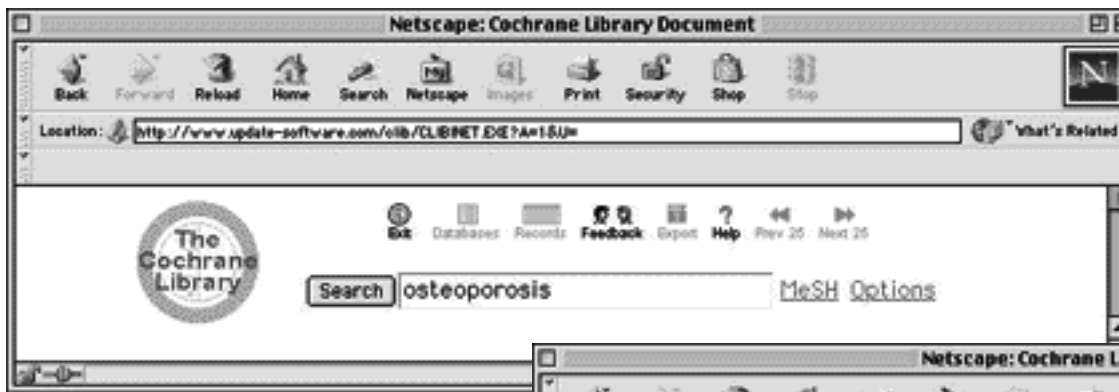
Det är viktigt att göra klart för sig centrala aspekter på frågeställningen innan sökningen startas (Ruta 1).

Först när vi besvarat dessa frågor har vi renodlat för oss själva vilken aspekt på den ursprungliga frågan som vi egentligen vill se besvarad, och hur den studie bör se ut som skall ge oss ett beslutsstöd.

Börja sökandet där det kliniska informationsvärdet är störst

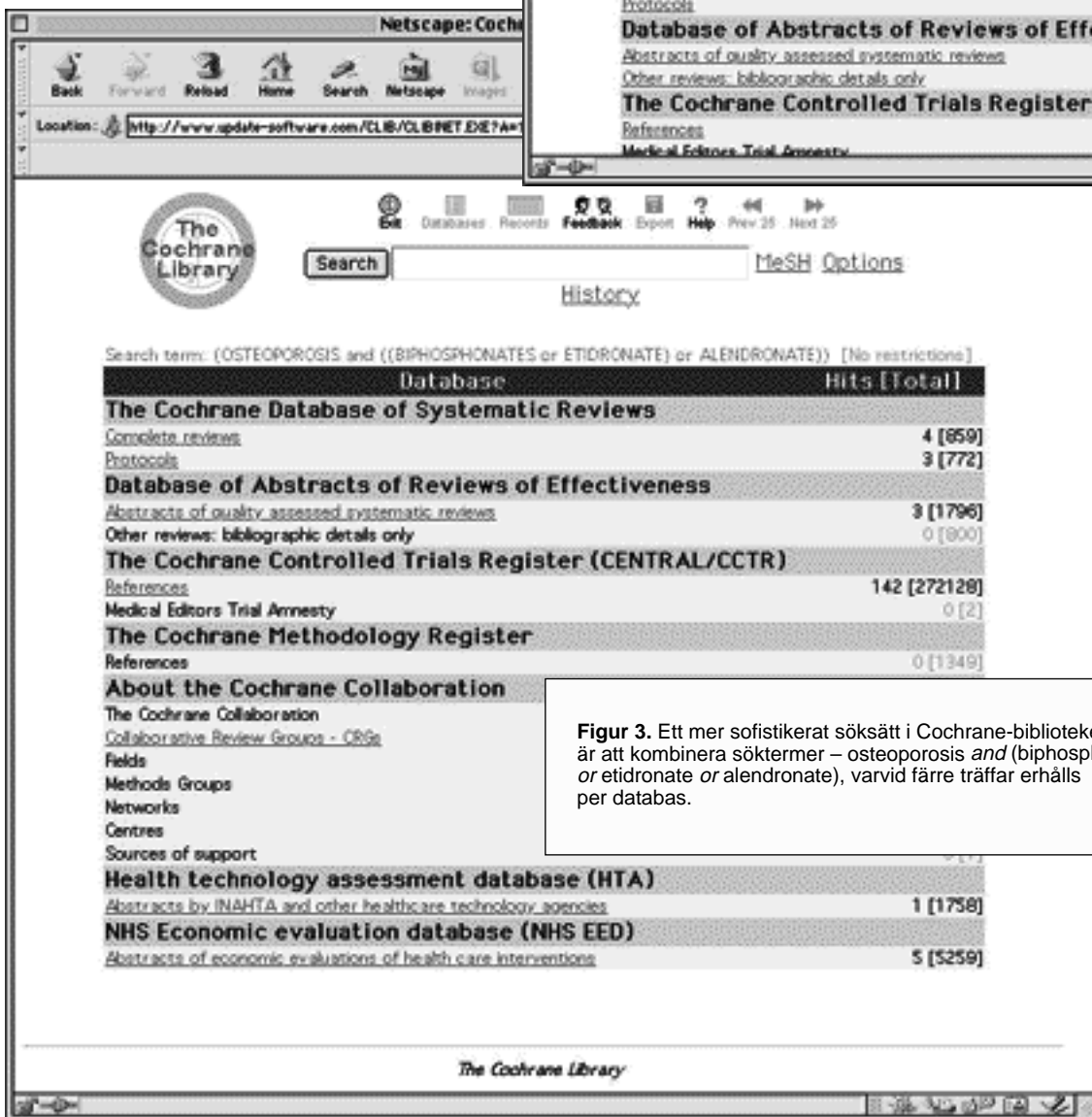
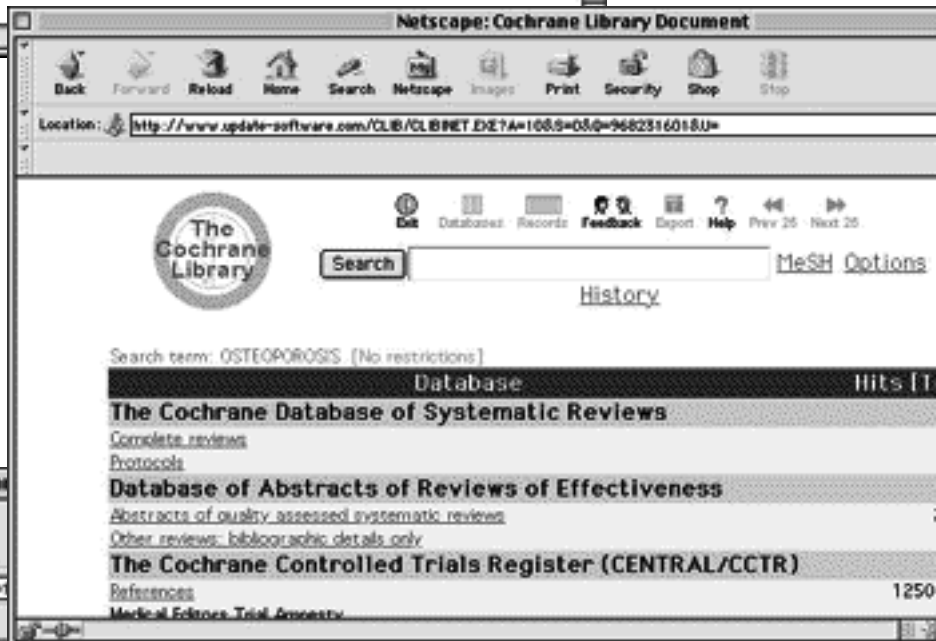
Nästa steg är att välja informationskälla. Vi börjar med att undersöka de källor som idag finns med för kliniker lättillgänglig, kritiskt granskad och kommenterad information, s k sekundärlitteratur. Vad vi mest av allt önskar är att hitta en bra systematisk översikt med fokus på vår problemställning [3].

- Vi börjar med *Clinical evidence*, ett terapikompendium under uppbyggnad som är baserat på evidensbaserad medicin och som uppdateras två gånger om

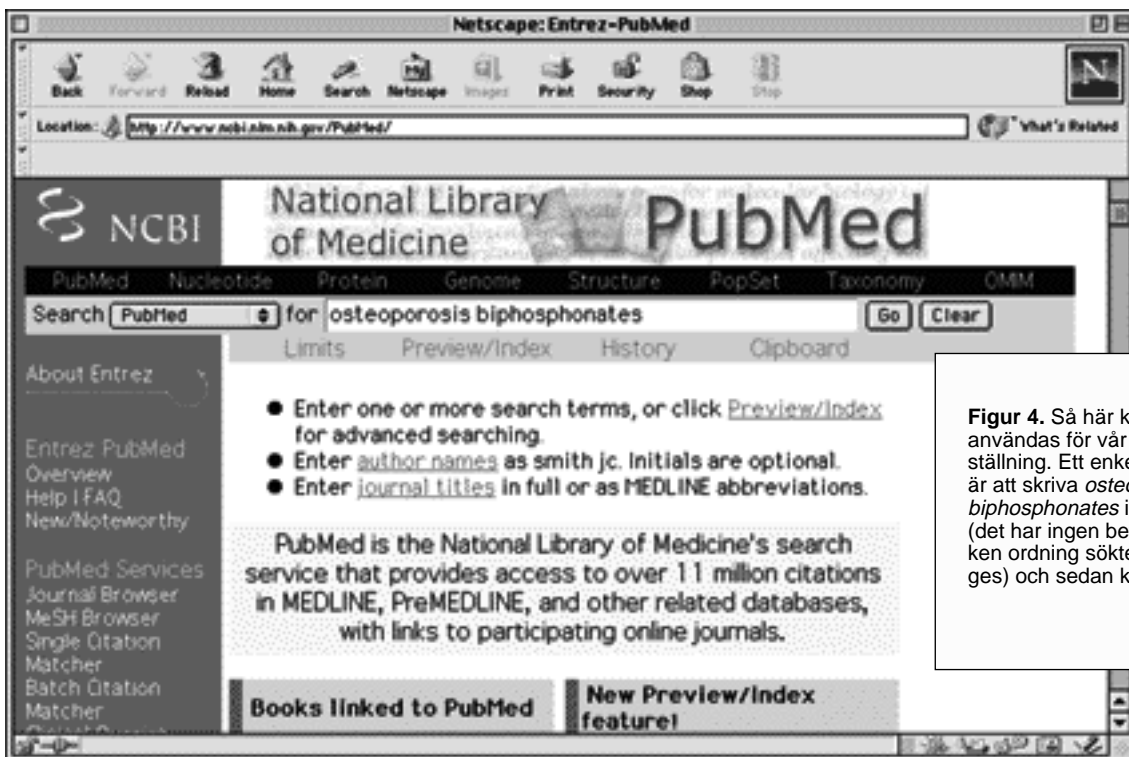


Figur 1. I sökrutan skrivs aktuellt sökord in. I detta fall vill vi ha svar på en frågeställning om osteoporos.

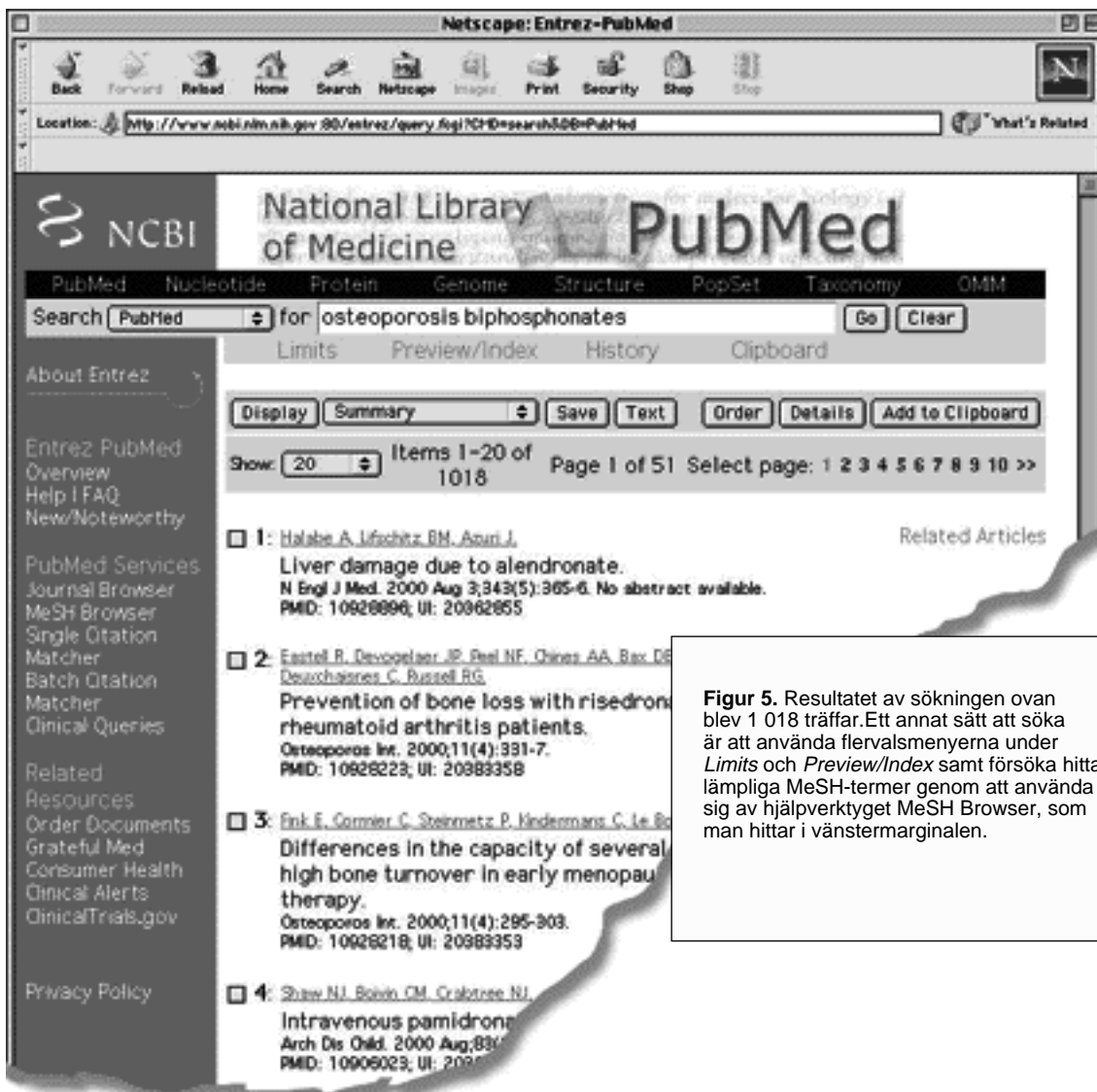
Figur 2. En av databaserna är Cochrane database of systematic reviews, som för närvarande innehåller 859 systematiska översikter. Dessutom finns projektbeskrivningar, »protocols», som beskriver det pågående arbetet med Cochrane-översikter. Dessa protokoll ger ofta bra nyckelreferenser och diskussion om metodproblem vid arbetet med den aktuella frågeställningen.



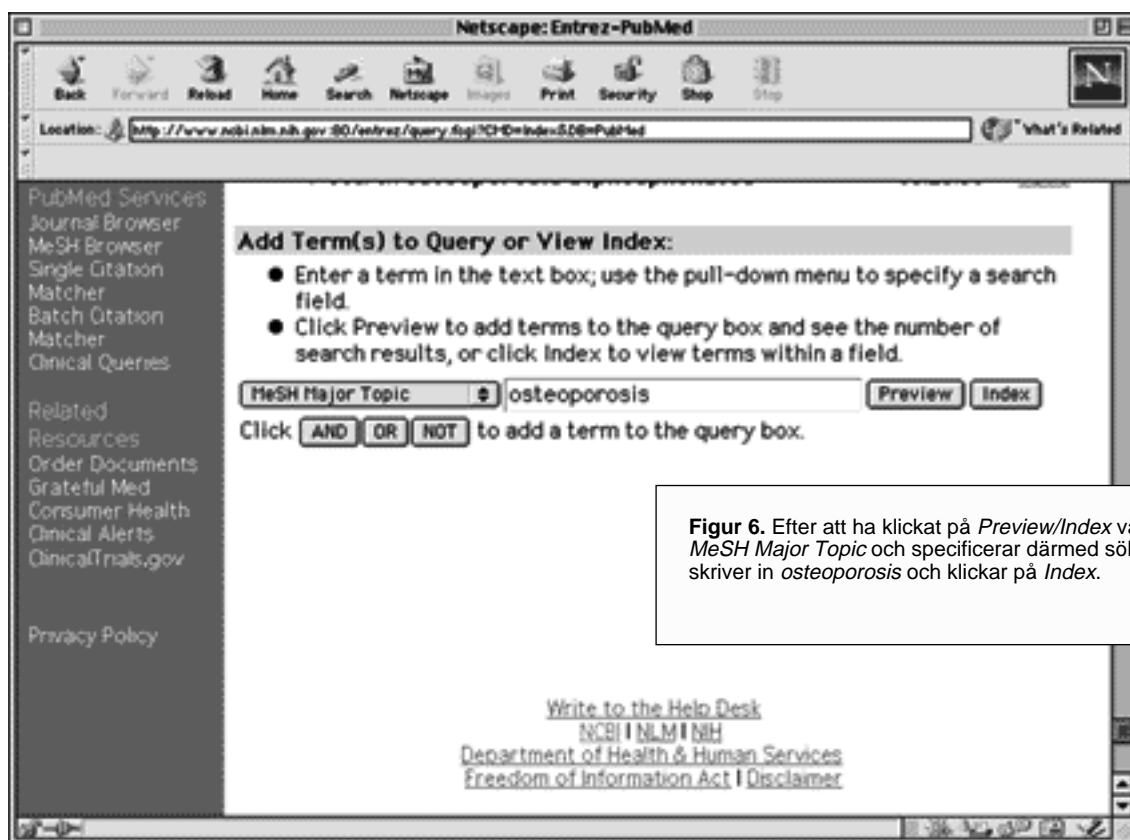
Figur 3. Ett mer sofistikerat söksätt i Cochrane-biblioteket är att kombinera söktermer – osteoporosis and (biphosphonates or etidronate or alendronate), varvid färre träffar erhålls per databas.



Figur 4. Så här kan PubMed användas för vår problemställning. Ett enkelt söksätt är att skriva *osteoporosis biphosphonates* i sökrutan (det har ingen betydelse i vilken ordning sökttermerna anges) och sedan klicka på Go.



Figur 5. Resultatet av sökningen ovan blev 1 018 träffar. Ett annat sätt att söka är att använda flervalsmenyerna under *Limits* och *Preview/Index* samt försöka hitta lämpliga MeSH-termer genom att använda sig av hjälpverktyget MeSH Browser, som man hittar i vänstermarginalen.



Figur 6. Efter att ha klickat på *Preview/Index* väljer vi *MeSH Major Topic* och specificerar därmed sökfält samt skriver in *osteoporosis* och klickar på *Index*.

året [4]. Sedan i somras finns det även i elektronisk form och sökbart på Internet. Den fullständiga versionen är lösenordsskyddad och bara tillgänglig för prenumeranter. Kompendiet är konstruerat för att besvara kliniska frågeställningar med väl underbyggda alternativ till åtgärder inom olika sjukdomsområden.

Denna källa innehöll dock ännu ingenting om behandling med bifosfonater vid osteoporos.

• Därefter söker vi i *Cochrane-biblioteket*, som innehåller ett antal databaser med systematiska litteraturöversikter över mer eller mindre breda kliniska frågeställningar [5]. Dessa databaser kan sägas vara en genväg till systematiskt kvalitetsgranskad information [6]. De är dessutom i förhållande till Medline volymmässigt betydligt mindre, och därmed lättare att söka i. En översikt över de databaser som ingår återges i Ruta 2.

I sökrutan för *Cochrane-biblioteket* skriver vi in *osteoporosis* (Figur 1), och söker därmed i alla de databaser som ingår. Som synes finns även möjlighet att använda MeSH-termer vid sökning (Ruta 3), men man bör då komma ihåg att inte all information i *Cochrane-biblioteket* är indexerad med MeSH-termer, varför man bör använda fri text plus MeSH. Under »Options» ges möjlighet bl a att specificera sökning till att gälla enbart texten i titlarna, eller

Ruta 1

1. Populationen eller patientgruppen:

Personer av båda könen oavsett ålder med en bentäthet som är så låg att den definitionsmässigt klassas som osteoporos, dock inte kortisoninducerad osteoporos, samt där risk för fraktur finns:

- med tidigare osteoporotisk fraktur;
- utan tidigare fraktur.

2. Interventionen:

Läkemedelsbehandling med något bifosfonatpreparat (alendronat eller etidronat).

3. Jämförelsen:

Vad skall kontrollgruppen ha behandlats med?

- placebo
- kalk och D-vitamin
- östrogen
- det andra bifosfonatpreparatet

4. Effektmåttet antal och typ av frakturer:

- symtomgivande
- samtliga frakturer

Skall studier som använt enbart bentäthet som effektmått accepteras?

Vilka biverkningar kan vi förvänta oss och vill vi se belysta?

Hur utfaller en hälsoekonomisk analys?

till enbart nya eller uppdaterade översikter.

• En av databaserna är *Cochrane database of systematic reviews*, som innehåller de systematiska översikter som *Cochrane-samarbetet* själv har producerat, för närvarande 859 stycken (Figur 2). Dessutom finns projektbeskrivningar, »protocols», som beskriver det pågående arbetet med *Cochrane-översikter*, vilka ofta ger bra nyckelreferenser och diskussion om metodproblem vid arbetet med den aktuella frågeställningen.

Dessa två *Cochrane-register* innehåller dokumenten i fulltext. Det innebär att vid vår sökning på osteoporosis fångas alla dokument in där termen finns någonstans i texten, såväl i titeln som i någon bisats. Antalet träffar blir ändå inte fler än att man relativt snabbt kan bläddra igenom titlarna genom att klicka på respektive understruket dataregister. För en utförlig beskrivning av arbets sättet vid sökning i *Cochrane-biblioteket* hänvisas till den svenska manual som finns på SBU:s webbsida.

Mer sofistikerade söksätt

Ett mer sofistikerat söksätt i *Cochrane-biblioteket* är att kombinera söktermer – *osteoporosis and (biphosphonates or etidronate or alendronate)*, varvid färre träffar erhålls per databas (Figur 3). Det visar sig att ingen av de fyra översikterna i »Complete reviews»

(Cochrane-samarbetets systematiska översikter) motsvarade vår problemställning. Två av översikterna handlade om kortisoninducerad osteoporos. Där emot visade sig två av protokollen, »protocols», handla om det vi sökte nämligen:

- Alendronate for osteoporosis in postmenopausal women (protocol); (färdig översikt slutet av 2000).
- Etidronate for osteoporosis in postmenopausal women (protocol); (färdig översikt 2001).

Något svar på vår problemställning får vi emellertid inte från dessa idag.

- Vi finner tre systematiska översikter i *Database of abstracts of reviews of effectiveness (DARE)* [7-9]. Dessa arbeten, som alla finns i tidskrifter som täcks av Medline, har genomgått en strikt kvalitetsgranskning av forskare vid universitetet i York, England.

I databasen DARE finns utförliga strukturerade sammanfattningar, »abstracts», med kommentarer, som bör kunna ge tillräckligt med information för att ge svar på vår problemställning (se »author's conclusions» i Ruta 4).

- Mycket färiska tidskriftsbaserade systematiska översikter kan vara idé att söka efter i *PubMed*, då det finns en viss eftersläpning i uppdateringen av Cochrane-biblioteket och DARE.

- Ytterligare ett intressant dokument hittar vi i databasen *Health technology assessment database (HTA)* [10]. HTA-databasen innehåller information om pågående och färdiga projekt från INAHTAs (internationellt nätverk för HTA-medicinska metodutvärderingar) medlemsorganisationer. Dokumentet är en rapport med en strukturerad sammanfattning tillgänglig i databasen (se »author's conclusions» i Ruta 5).

Där finns även en Internet-adress angiven, från vilken det går att ladda ner rapporten i sin helhet, få en sammanfattning av dokumentet samt möjlighet att beställa den tryckta rapporten.

- Av övriga databaser som ingår i Cochrane-biblioteket är *Cochrane Controlled Trials Register (CENTRAL/CCTR)* värd att nämna. Den innehåller alla kontrollerade studier (framför allt randomiserade) som har samlats in av Cochrane-samarbetet genom sökning i elektroniska källor samt genom manuell sökning i tidskrifter.

För närvarande finns 270 000 studier i databasen innehållande bearbetade uppgifter i form av referenser, oftast med abstract och nyckelord. Endast kliniska humanförsök ingår.

Ruta 2

Databaser i Cochrane-biblioteket:

- The Cochrane Database of Systematic Reviews (CDSR)
- Database of Abstracts of Reviews of Effectiveness (DARE)
- The Cochrane Controlled Trials Register (CENTRAL/CCTR)
- Health Technology Assessment database (HTA)
- NHS Economic Evaluation Database (NHSEED)

Ruta 3

MeSH-systemet

MeSH (Medical Subject Headings) är de nyckelord som används för att indexera/klassificera (beskriva) dokumenten i Medline och i många andra medicinska databaser. Det finns idag omkring 20 000 MeSH-termer i den tesaurus som NLM (US National Library of Medicine) har sammanställt och kontinuerligt håller uppdaterad.

- En annan databas i Cochrane-biblioteket, som kan vara av intresse att känna till, är *NHS Economic evaluation database (NHSEED)*. Den innehåller analyserade och kommenterade kostnad-effektivitetsstudier. Det fanns fem titlar med relevans för vår problemställning [11-15].

Vad göra när systematiska översikter saknas?

Om inte något acceptabelt svar har erhållits i föregående källor är nästa steg att söka *Medline*, en av de viktigaste medicinska databaserna. Medline finns sökbar på olika system – t ex Ovid, Silver Platter, Grateful Med. Ett av de mest populära idag är PubMed, som är gratis och innehåller Medline och PreMedline. I PreMedline ingår färskt material från ett urval av de tidskrifter som täcks av Medline och som ännu inte indexerats, dvs är utan nyckelord (MeSH-termer).

En stor fördel med att använda MeSH-termer som sökord är att man slipper använda synonyma termer för ett begrepp (*neoplasms* täcker cancer, tumor, carcinoma etc), och man behöver inte tänka på olika stavningsvarian-

ter (estrogen eller oestrogen, meta-analysis eller metaanalysis).

Trots att vi hittade en del i föregående källor om osteoporos och bifosfonater vill vi ändå visa hur PubMed kan användas för vår problemställning. Ett enkelt söksätt är att skriva *osteoporosis biphosphonates* i sökrutan (Figur 4) (det har ingen betydelse i vilken ordning söktermerna anges) och sedan klicka på *GO*. Resultatet blev 1 018 träffar (Figur 5).

Detaljerad information om hur sökningen har gått till får man genom att klicka på *Details*. Det visar sig då att programmet har letat fram lämpliga MeSH-termer samt sökt på ord i titel och abstract (textord). För att gå vidare och begränsa sökningen till att gälla enbart en viss studiedesign, t ex metaanalyser, lägger man till *meta-analysis* i sökrutan, och antalet träffar minskar drastiskt till sju.

Ett annat sätt att söka är att använda flervalsmenyerna under *Limits* och *Preview/Index* (Figur 4 och 5) samt försöka hitta lämpliga MeSH-termer genom att använda sig av hjälpverktyget MeSH Browser, som man hittar i vänstermarginalen.

För att visa vad detta alternativa söksätt innebär klickar vi på *Preview/Index*, väljer *MeSH Major Topic* och specificerar därmed sökfält, skriver in *osteoporosis* och klickar på *Index* (Figur 6). Genom att titta igenom listan av sökbara termer upptäcker vi att det även är möjligt att välja en s k »subheading» kopplad till MeSH-terminen, t ex *osteoporosis/drug therapy* och *osteoporosis/prevention and control*. Vi väljer dessa båda termer samt klickar på *AND*. Därmed skrivs dessa söktermer in i sökrutan.

Därefter gör vi samma sak med *diphosphonates* men väljer *MeSH Terms* som sökfält. Genom MeSH Browser har vi fått veta att den korrekta MeSH-terminen är *diphosphonates* samt att *alendronate* och *etidronate disodium* finns som underordnade termer. Det innebär att alla dessa termer söks, alltså även underordnade termer. Detsamma gäller vid sökning på *osteoporosis* – även den underordnade MeSH-terminen *osteoporosis, postmenopausal* söks automatiskt i PubMed. Med kombinationen av söktermerna *osteoporosis / drug therapy OR osteoporosis / prevention & control AND diphosphonates* erhålls 544 träffar.

Det finns möjligheter att ställa ytterligare villkor, förutom det rent ämnesmässiga. Det kan gälla t ex publikationstyp eller studiedesign, publiceringsperiod (i stället för att söka hela perioden från och med år 1966 som ingår i Medline), språk, ålder och kön. Alla dessa möjligheter finns under rubriken

Limits. Här kan vi välja *ex randomized controlled trial* eller *meta-analysis* som publikationstyp. *Systematic review* finns inte som specifik indexerings-term utan går ofta under *meta-analysis* eller måste sökas som fri text. Vi väljer *meta-analysis*, och antalet träffar minskar till sex.

Alla dessa sex referenser finns inte med bland de sju som erhöles i det första sökexemplet i PubMed. I det senare är alla dokumenten indexerade med *meta-analysis* som publikationstyp. I det förra fanns tre dokument med ordet *metaanalysis* nämnt enbart i titel eller abstract. Detta visar att det finns många olika söksätt i samma databas och med hjälp av samma söksystem, och att en sökning ofta kan resultera i något olika resultat. Det finns ingen »golden standard». Det facit man har är utfallet, dvs om de referenser jag hittar motsvarar det jag söker efter, alltså ger svaret på min fråga. Om inte, får man modifiera sin sökstrategi med någon av alla de andra möjligheter som finns till hands. Här har enbart visats några få sätt.

PubMed-funktioner som kan förenkla

I syfte att underlätta för kliniker att snabbt och enkelt hitta relevant information i PubMed har ett speciellt gränssnitt tagits fram, *Clinical Queries* (se vänsterkanten på Figur 6). Vi testar på vår problemställning och skriver *osteoporosis biphosphonates* i sökrutan samt prickar för *therapy* respektive *specificity*. Resultatet blir 151 träffar, vilket i det här fallet kanske är lite för många för en tidspressad kliniker att bläddra igenom. Det visar sig vara både randomiserade kontrollerade studier och studier med mindre strikt studiedesign, där dubbelblind eller placebo nämns någonstans i referensen.

PubMed innehåller omkring 11 miljoner referenser och ökar kontinuerligt allteftersom nytt material publiceras. Detta gör att olika möjligheter behövs för att hitta rätta söktermer, kombinera dessa och begränsa sökningen på olika sätt för att kunna filtrera fram det som motsvarar en frågeställning. En omtyckt finess i PubMed är att kunna klicka på *Related articles* i anslutning till en referenstitel. Det innebär att sökprogrammet tar fram ytterligare referenser, som innehåller ungefär samma kombination av MeSH-termer. Det är ett sätt att genom repetitiva processer fokusera alltmer.

Förutom hjälpinstruktionerna, som finns tillgängliga i respektive söksystem, finns en träningsmanual för PubMed på Internet. För den som vill avsätta tid för att lära sig mer med hjälp av handledda övningar vid dator finns bl a Karolinska institutets biblioteks halv-

Ruta 4

Author's conclusions:

The results of the overview confirm that both calcitonin and etidronate increase BMD in post menopausal women. The existing evidence does not allow the superiority of either of the 2 drugs to be established. Pooled results suggest that calcitonin prevents more vertebral fractures but the clinical trials are very heterogeneous. Etidronate seems to achieve a greater increase in vertebral BMD but the evidence still suffers from lack of homogeneity. Neither of the 2 drugs provide evidence of efficacy in the prevention of hip fractures. [7]

In post menopausal women with osteoporosis, treatment with alendronate reduces the risk of nonvertebral fracture over at least three years. [8]

Only alendronate and vitamin D plus calcium have clearly demonstrated a fracture benefit with alendronate providing the greatest relative risk reduction. Quality clinical trial fracture data for calcitonin, etidronate, fluoride, hormone replacement therapy, parathyroid hormone, calcitriol (and other vitamin D preparations), vitamin D and calcium monotherapy and selective oestrogen modulators are either lacking or inconclusive or published only as abstracts. [9]

Ruta 5

Author's conclusions:

The authors find that the choice of target population is critical to the cost-utility of bisphosphonate treatment. If treatment is targeted towards individuals at high risk of fracture then both alendronate and etidronate appear to be reasonably cost-effective. However they cannot support the use of etidronate, as RCTs have not detected a reduction in non-vertebral fracture rate, despite suggestions from proxy measures that this may result. Alendronate is preferred, despite it being a more expensive drug, as the quality of evidence for its effectiveness is higher. [10]

dags- och heldagskurser. Många andra biomedicinska bibliotek ordnar liknande utbildningar.

»One-stop shopping»

Medline, som produceras av US National Library of Medicine, täcker omkring 4 000 vetenskapliga tidskrifter, vilket utgör mindre än 20 procent av det totala medicinska tidskriftsmaterial som finns i världen, dock lär det vara det viktigaste. Arbete pågår vid US National Institutes of Health för att skapa ett globalt, sökbar och sammanlänkat elektroniskt arkiv, »a one-stop shop», för vetenskapligt material inom inte bara medicin utan också naturvetenskap generellt [16].

Andra initiativ av typen »One-stop shop» finns, bl a det medicinska sökverktyget SUMSearch, som för en frågeställning väljer ut och söker i flera medicinska källor (Merck Manual, Medline, DARE, National Guideline Clearinghouse, Bedside Diagnosis, Food and Drug Administration). Vilka som söks beror på typ av frågeställning samt utfallet i några av databaserna.

Vid sökning på vår problemställning om osteoporos och bifosfonater erhöles efter flera minuters väntan, dock med möjlighet att läsa senaste nytt från ReutersHealth under tiden, svar i form av antalet träffar uppdelade på typ av dokument samt databas. Så gott som alla träffar (72 av 73) var PubMed-referenser.

Ett annat sökverktyg är det som används för TRIP-databasen, som är en sammanlänkning av 26 mer eller mindre evidensbaserade Internet-källor. Vid sökning på osteoporos erhöles 59 träffar, av vilka många härrörde från de databaser som redan beskrivits.

Internationellt efterfrågas och produceras allt fler kliniska riktlinjer, »guidelines», och kravet idag är att dessa skall vara evidensbaserade. Vid sökning på frågeställningen om osteoporos och bifosfonater i den amerikanska guideline-databasen National Guideline Clearinghouse erhöles sex träffar, som alla finns tillgängliga på Internet i fulltext eller i form av mycket utförliga strukturerade sammanfattningar.

Ett sätt att hitta genvägar till kvalitetsgranskad information

Varje klinisk problemställning är unik och fordrar sin speciella strategi beroende på omständigheterna. Med denna exposé har vi försökt visa några olika möjligheter att gå tillväga, och inte minst påminna om de genvägar som finns till redan kvalitetsgranskad och kommenterad information. De möjligheter som står till buds ökar för varje dag och det är omöjligt för den kliniskt verksamme att hålla sig a jour med alla

alternativ och vad som är bäst för stunden.

Med databaserna i Cochrane-biblioteket och med *PubMed* samt *Clinical evidence* är man ganska väl försörjd i informationshänseende i den vardagliga sjukvården. Att snabbt få ett svar på en frågeställning är dock inte alltid möjligt, även om *Clinical evidence* har den målsättningen.

För övrigt gäller det att kunna tolka vad som står i strukturerade sammanfattningar eller i guidelines, och applicera denna kunskap i det kliniska arbetet. Det är en fördel att ha dokument tillgängliga i fulltext, som Cochranes systematiska översikter. Fler och fler medicinska tidskrifter läggs dock ut i sin helhet elektroniskt, dock ofta via lösenord. Det gäller att hålla sig informerad om vad sjukhusbiblioteket eller tillhörande forskningsbibliotek kan erbjuda i det avseendet. För kliniker som tillhör Karolinska institutet finns idag mer än 1 200 tidskrifter tillgängliga elektroniskt i fulltext av bibliotekets totalt omkring 2 700 tidskrifter.

Referenser

1. Werkö L. Evidensbaserad medicin. Ett konkret redskap för vårdens verkstadsgolv. *Läkartidningen* 2000; 97: 2710-3.
2. Eliasson M. Evidensbaserad medicin »begins at home». Incitament för bättre beslutsunderlag finns i den kliniska vardagen. *Läkartidningen* 2000; 97: 3860-5.
3. Eliasson M. Den systematiska översikten grundval i evidensbaserad medicin. Ett av de viktigaste tillskotten till den kliniska medicinen under senaste årtiondet. *Läkartidningen* 2000; 97: 2726-8.
4. Clinical Evidence – a compendium of the best available evidence for effective health care. BMJ Publishing Group, Issue 3, June 2000.
5. Gøtzsche P. Cochranesamarbetets betydelse för evidensbaserad medicin. *Läkartidningen* 2000; 97: 3180-2.
6. Eliasson M. En guide till kritisk läsning av kliniska artiklar. *Läkartidningen* 2000; 97: 3178-9.
7. Cardona JM, Pastor E. Calcitonin versus etidronate for the treatment of postmenopausal osteoporosis: a meta-analysis of published clinical trials. *Osteoporos Int* 1997; 7:165-74.
8. Karpf DB, Shapiro DR, Seeman E, Ensrud KE, Johnston CC, Adami S et al. Prevention of nonvertebral fractures by alendronate: a meta-analysis. *JAMA* 1997; 277(14): 1159-64.
9. Meunier PJ. Evidence-based medicine and osteoporosis: a comparison of fracture risk reduction data from osteoporosis randomized clinical trials. *International Journal of Clinical Practice* 1999; 53(2): 122-9.
10. Best L, Milne R. Bisphosphonates (alendronate and etidronate) in the management of osteoporosis. Southampton: Wessex Institute for Health Research and Development (DEC Report No. 79), 1998.
11. Francis RM, Anderson FH, Torgerson DJ. A comparison of the effectiveness and cost of treatment for vertebral fractures in women. *Br J Rheumatol* 1995; 34(12): 1167-71.
12. Rosner AJ, Grima DT, Torrance GW, Bradley C, Adachi JD, Seibald RJ et al. Cost ef-

fectiveness of multi-therapy treatment strategies in the prevention of vertebral fractures in postmenopausal women with osteoporosis. *Pharmacoeconomics* 1998; 14(5): 559-73.

13. Visentin P, Ciravegna R, Corcelli F, Fabris F. Cost-effectiveness of hip fracture prevention. *Epidemiol Prev* 1998; 22(1): 44-8.
14. Reid DM, Torgerson DJ. Pharmacoeconomic aspects of intermittent cyclical etidronate therapy in the treatment of postmenopausal and corticosteroid-induced osteoporosis. *Reviews in Contemporary Pharmacotherapy* 1998; 9(4): 287-92.
15. Ankjaer-Jensen A, Johnell O. Prevention of osteoporosis: cost-effectiveness of different pharmaceutical treatments. *Osteoporos Int* 1996; 6(4): 265-75.
16. Rabow I. Elektronisk publicering både problem och möjligheter. Forskarvärldens ansvar att driva utvecklingen åt rätt håll. *Läkartidningen* 2000; 97: 2091-5.

*

Några nyttiga Web-adresser

Clinical Evidence:

<http://www.clinicalevidence.org>

Cochrane Library:

<http://www.update-software.com/clibhome/clib.htm> (lösenord erfordras) eller <http://www.slfl.se> gå till Doktors sida (gratis tillgänglig för medlemmar i Läkarförbundet)

Cochrane Collaboration

(enbart »abstract» av Cochrane-översikter): <http://www.update-software.com/ccweb/cochrane/revabstr/mainindex.htm> (utan lösenord)

SBU:

<http://www.sbu.se>

PubMed (Medline):

<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=PubMed>

Träningsmanual för PubMed på Internet:

http://www.nlm.nih.gov/pubs/web_based.html

Karolinska institutets biblioteks kurser:

http://www.kib.ki.se/edu/bib_kurs/

SUMSearch:

<http://sumsearch.uthscsa.edu/searchform4.htm>

TRIP (Turning Research Into Practice):

<http://www.ceres.uwcm.ac.uk/frameset.cfm?section=trip>

National Guideline Clearinghouse:

<http://guideline.gov>

Summary

How to search for clinical answers on the Internet

Viveka Alton, Mats Eliasson

Läkartidningen 2000; 97: 4426-34

To achieve an evidence-based clinical practice it is vital to be able to rapidly access valid, relevant and updated information. Systematic reviews, HTA-reports and health economic evaluations are most easily found in the Cochrane Library, available by subscription on the web. The continually updated »Clinical Evidence» covers many clinical fields. A search on Medline will mostly yield primary research but also systematic reviews. Search-engines such as SUMSearch and TRIP could be useful by scanning multiple databases relevant to evidence-based medicine.

Correspondence: Viveka Alton, SBU, Box 5650, SE-114 86 Stockholm, Sweden.

E-mail: alton@sbu.se

Särtryck

LÄKARTIDNINGEN

När konsensus saknas om hur läkaren bör behandla, spelar den beprövade erfarenheten stor roll. Det 48-sidiga häftet innehåller 32 korta, praktiskt inriktade artiklar med anknytning till vårdens vardag och vänder sig till alla kliniskt verksamma läkare. Förutom diagnostik med terapi speglas goda exempel på prevention, ledningsfrågor och administration.

Priset är 55 kr.

Enligt min erfarenhet



Beställer härmed.....ex av "Enligt min erfarenhet"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till **LÄKARTIDNINGEN**
Box 5603
114 86 Stockholm

Faxnummer: **08-20 74 35**

www.lakartidningen.se
under särtryck, böcker

Evidensbaserad medicin

Så graderas en studies vetenskapliga bevisvärde och slutsatsernas styrka

Studiers bevisvärde

Högt bevisvärde

Tillräckligt stor studie, lämplig studietyp, väl genomförd och analyserad. Kan vara en stor, randomiserad kontrollerad studie (RCT) när det gäller utvärdering av en behandlingsform. För övriga områden: Uppfyller väl på förhand uppställda kriterier.

Medelhögt bevisvärde

Behandlingseffekter: Kan vara stora studier med kontroller från andra geografiska områden, matchade grupper eller liknande. För övriga områden: Uppfyller delvis på förhand uppställda kriterier.

Lågt bevisvärde

Skall ej ligga som enda grund för slutsatser, t ex studier med selekterade kontroller (retrospektiv jämförelse mellan patientgrupper som fått respektive inte fått en viss behandling), stora bortfall eller andra osäkerheter. För övriga områden: Uppfyller dåligt på förhand uppställda kriterier.

Grundvalen för evidensbaserad medicin är den systematiska översikten. Här används en strikt metodik för att finna, granska och kvalitetsgradera relevanta studier vars resultat sammanfattas, kvantitativt när så är lämpligt och möjligt, i form av en metaanalys. Detta har beskrivits närmare av Mats Eliasson i *Läkartidningen* 22/00 [1]. Den springande punkten är

att undvika systematisk feltolkning, bias. Förutom strikta metoder, som fastställts före starten, krävs adekvat kompetens av granskarna, gärna medverkan av såväl förespråkare som kritiker, beaktande av intressekonflikter, gender och var man bor.

När det gäller krav på kvalitet för att en studies slutsatser skall kunna tillämpas får man ofta definiera en miniminivå. Det kan innebära att man går förbi studier som helt saknar kontrollgrupper. Kanske vill man inte heller beakta studier som innehåller blandade patientgrupper, studier med otillräcklig uppföljningstid, eller sådana där man inte har några mått på resultat av betydelse för patientens välbefinnande utan enbart s k »surrogatvariabler» som t ex laboratorievärden.

Studiers bevisvärde

De studier som ligger över miniminivån kan indelas i olika grad av bevisvärde: från högt till medelhögt och lågt (se separat ruta). Detta avspeglar studiernas pålitlighet när det gäller att besvara den aktuella frågeställningen. Studier med högt bevisvärde avseende olika behandlingsformer bör vara stora, väl genomförda, randomiserade kontrollerade studier.

När det gäller diagnostik, riskbedömning och prognos får man titta på andra typer av studier vilkas bevisvärde också kan graderas från högt till lågt för den fråga som är aktuell.

Några viktiga krav för bedömning av en studies bevisvärde presenteras i separat ruta på motstående sida [2]. Den kan användas som en checklista för att gradera ett arbetes metodologiska kva-

Författare

MONA BRITTON

professor, medicinskt sakkunnig, Statens beredning för medicinsk utvärdering (SBU), Stockholm.

litet. Ju fler kriterier som är uppfyllda desto bättre. Det går dock inte att ange några absoluta gränser. Trots tabellens omfattning går det inte att få en fullständig förteckning som utesluter alla frågetecken, har helt klara gränser och blir odiskutabel. Till sist måste bedömningar göras utifrån ämneskunskap, vetenskaplig kompetens och sunt omdöme.

Slutsatsers evidensstyrka

Styrkan hos det vetenskapliga underlaget för en slutsats baserad på alla utvalda studier benämns i engelsk litteratur »level of evidence», evidensstyrka. Det kan vara av värde för läsare av systematiska litteraturöversikter att veta hur säker en slutsats är. Detta belyser också risken för hur snabbt den kan ändras. Vid starkt vetenskapligt stöd bör några mindre studier inte kunna påverka slutsatserna nämnvärt, medan man vid svagare underlag måste vara observant på det som tillkommer.

För att gradera evidensstyrka måste alla accepterade studier vara kritiskt granskade, deras bevisvärde bedömt och fakta extraherade. Först då är det dags att göra en summering och sedan dra slutsatser från den. Slutsatserna kan visa sig ha olika solid vetenskaplig grund (Tabell I).

Om flera stora studier, från olika centra och med en för frågan lämplig

Tabell I. Gradering av evidensstyrka för slutsatser.

1: Starkt vetenskapligt underlag	2: Måttligt starkt vetenskapligt underlag	3: Begränsat vetenskapligt underlag	4: Otillräckligt vetenskapligt underlag
Minst två studier med högt bevisvärde eller god systematisk översikt	En studie med högt bevisvärde plus minst två med medelhögt bevisvärde	Minst två studier med medelhögt bevisvärde	Annat underlag: Vad?

Inklusionskriterier

(baserad på en tabell i Tidsskrift for Den norske Lægeforening [2])

For systematiske oversikter:

Oversikten må inneholde en metodeseksjon som beskriver inklusionskriteriene og søkestrategien

For studier av effekt:

Fordelingen av pasienter til behandlingene må være randomisert

For studier av nøyaktighet av diagnostiske tester:

Det må være en uavhengig, blind sammenlikning med en referansestandard (gullstandard) for diagnosen
Den diagnostiske testen må ha blitt evaluert i et relevant spekter av pasienter (tilsvarende den gruppen som testen vil bli brukt på i praksis)

For studier av skade/bivirkninger:

Det må være en klart identifisert sammenlikningsgruppe for dem som har økt risiko for, eller som har det utfallet som er av interesse (enten fra randomiserte, kvasirandomiserte eller ikke-randomiserte, kontrollerte forsøk, kohortstudier med statistiske justeringer for å skape sammenliknbare grupper eller case control-studier)
Eksponering for intervensjonen og de kliniske utfall må være målt på samme måten i begge grupper
Oppfølgingstiden må være lang nok

For studier av insidens/prevalens:

Det må være et representativt (tilfeldig, konsekutivt eller totalt) utvalg
Det må være klart definerte diagnostiske kriterier

Videre vurdering

kun av inkluderte studier

For systematiske oversikter:

Var kriterier for vurdering av validitet av inkluderte studier og analysene beskrevet i metodeseksjonen?

Var det inkludert ≥ 1 studie som møter inklusionskriteriene for enkeltstudier?

Var resultatene av de inkluderte studiene som er relevante for anbefalingene i hovedsak konsistente?

For studier av effekt:

Var randomiseringsprosessen lukket?

Var sammenlikningsgruppene like i utgangspunktet?

Var $\geq 80\%$ av pasientene som ble inkludert i forsøket redegjort for i konklusjonen?

Var studien dobbeltblindet?

For studier av nøyaktighet av diagnostiske tester:

Var referansestandard applisert uavhengig av resultatet

av den diagnostiske testen?

Hvis studien gjaldt et diagnostisk kriterie-sett eller algoritme, ble dette validert i en annen kohort av pasienter?

For studier av skade/bivirkninger:

Var sammenlikningsgruppene like på andre viktige måter enn eksponeringen for intervensjonen?

Var $\geq 80\%$ av pasientene som ble inkludert i studien redegjort for i konklusjonen?

Var det brukt blindet vurdering eller objektive mål for å måle eksponering eller utfall?

For studier om insidens eller prevalens:

Var de diagnostiske kriterier applisert på samme måten til hele utvalget slik at det var liten risiko for falskt positive eller falskt negative tilfeller?

design og høgt bevisvärde, givitt samma resultat blir det vetenskapliga underlaget starkt och slutsatsen pålitlig (evidensstyrka 1). Det gäller även då det finns en god systematisk översikt av flera väljorda studier.

Om endast en studie med högt bevisvärde finns, men resultaten stöds av flera studier med medelhögt bevisvärde, blir evidensstyrkan måttligt stark (evidensstyrka 2).

Begränsad styrka (evidensstyrka 3) föreligger när ingen studie med högt bevisvärde men minst två med medelhögt bevisvärde finns.

Under denna nivå är det vetenskapliga underlaget otillräckligt; man måste ange vad eventuella slutsatser i så fall baserar sig på (evidensstyrka 4).

Förutsättningen för det angivna är att studierna pekar i samma riktning och att inget talar emot slutsatsen. Vid mindre divergenser kan styrkegraden sänkas.

SBU-rapporterna anger evidensstyrka

I fortsättningen skall vi på SBU ange evidensstyrka i rapporternas slutsatser. Det har vi också gjort i de senaste rapporterna i form av bokstavsbeaktningar. Bokstäverna A–D har använts i stället för siffrorna 1–4. Nuvarande ändring baserar sig på en önskan om att skilja graderingen på evidensstyrka från den vid riktlinjer, rekommendationer, »guidelines».

Internationellt har man ofta använt siffror (arabiska eller romerska) för styrkan i evidensen, bokstäver för styrkan i rekommendationer [2-5]. Vi anpassar oss nu till det bruket.

Många andra hänsyn

Vid utarbetande av guidelines bör evidens vara en viktig grundval. Därutöver måste man ta hänsyn till de prioriteringar som krävs av hälso- och sjukvårdslagen. Även kostnadseffektiviteten behöver beaktas samt möjligheterna att genomföra de evidensbaserade slutsatserna i praxis. Här kan lokala för-

hållanden göra att man behöver modifiera rekommendationerna.

Sådana aspekter måste vägas in när vårdprogram och riktlinjer upprättas. Därför är styrkan på evidens avseende effekter och biverkningar en viktig grundläggande faktor, dock inte den enda.

Referenser

1. Eliasson M. Den systematiske översikten grundval i evidensbaserad medicin. Läkartidningen 2000; 97:2726-28.
2. Oxman AD, Flottorp S, Cooper JG. Nivået på dokumentationen og styrken av anbefalingene i kliniske retningslinjer. Tidsskr Nor Lægeforen 2000; 120: 1791-5.
3. National Health Service (NHS). Levels of evidence and grades of recommendations. <http://ceb.m.jr2.ox.ac.uk/docs/levels.html>
4. Grading system for recommendations in evidence-based clinical guidelines. Notes and checklists. Scottish Intercollegiate Guidelines Network (SIGN), September 1999.
5. Woolf SH, DiGuseppi CG, Atkins D, Kamerow DB. Developing evidence-based clinical practice guidelines: Lessons learned by the US Preventive Services Task Force. Annu Rev Public Health 1996; 17: 511-38.

Är artikeln värd att läsa?

Praktisk handledning i konsten
att kritiskt granska en randomiserad behandlingsstudie

Det finns många anledningar att läsa behandlingsstudier. Man kan ögna igenom studier i sin egen specialisttidskrift för att få en översiktlig uppfattning om vad som händer: är nya spännande behandlingar på väg, eller har ytterligare data publicerats angående interventioner man själv arbetar med i sin vardag? Med god tillgänglighet till elektroniska databaser såsom Cochrane-biblioteket och Medline via Internet blir det allt vanligare att klinikern själv aktivt söker för att besvara en specifik klinisk fråga. Ibland är detta behov utgående från ett enskilt patientfall, men ofta kan det finnas ett uppdrag att ta fram riktlinjer eller underlag för vårdcentralen, kliniken eller läkemedelskommittén. För kliniska forskare finns ofta ett behov att fördjupa sökning och granskning för att kunna jämföra egna fynd med tidigare publikationer, sammanställa en systematisk översikt eller genomföra en medicinsk metodvärdering, exempelvis en rapport från Statens beredning för medicinsk utvärdering (SBU).

Självfallet kommer kraven på det kritiska läsandet att variera i dessa olika situationer, men vissa grundelement är gemensamma och lätta att ta till sig. Fokus i denna artikel är att ge den kliniskt verksamme läkaren, utan egen forskningserfarenhet, några enkla regler och råd för att i sin vardag bättre och snabbare kunna urskilja de bra studierna. Sådana bör läsas noggrannare och i vissa fall ligga till grund för förändringar i det vardagliga kliniska handlandet.

Leta kostnadseffektivt efter bästa kunskapsunderlaget

Om utgångspunkten är ett kliniskt problem bör din kliniska fråga vara väl

Författare

MATS ELIASSON

med dr, överläkare, Sunderby sjukhus, Luleå; projektsamordnare, Statens beredning för medicinsk utvärdering (SBU); medicinsk redaktör (evidensbaserad medicin), Läkartidningen.

E-post: mats.eliasson@nll.se

fokuserad. Strukturera och notera faktorer såsom sjukdomens karaktär, patientgruppens speciella karakteristika, behandlingens detaljer, kontrollgruppens behandling, vilka patientrelaterade effektmått som är viktiga och vilka potentiella biverkningar som bör vara belysta.

Kort sagt, du bör skissa på den optimala studien för att besvara din frågeställning [1]. Genom en sökning, gärna baserad på MeSH-termer, i Cochrane-biblioteket och PubMed identifierar du systematiska översikter och enskilda behandlingsstudier [2]. Behandlingsstudierna kan antingen vara experimentella, dvs randomiserade och kontrollerade studier (RCT), eller observationsstudier, dvs kohortstudier och fall-kontrollstudier.

Denna artikel kommer helt att handla om de randomiserade studierna, eftersom de har störst möjlighet att tillförlitligt besvara frågan huruvida en behandling leder till nytta för patienten [3, 4].

Påtagligt många frågor om behandling kan inte och kommer inte att kunna besvaras med randomiserade studier. Det är därför väsentligt att inte avfärda andra forskningsmetoder. Den evidensbaserade medicinens ursprungligt enögda fokus på RCT håller nu på att kompletteras med en intensiv debatt om hur observationsstudier och kvalitativa ansatser kan kritiskt granskas och användas som beslutsunderlag för behandlingsrekommendationer. Läkartidningen kommer att lyfta fram denna diskussion i kommande nummer.

Om du hittar en systematisk översikt som berör din fråga har du en bra utgångspunkt, eftersom mycket av ditt arbete redan är utfört och systematiserat. Kom dock ihåg att även systematiska översikter, med eller utan metaanalyser, måste granskas kritiskt [5]. I det fall du inte kan identifiera någon systematisk översikt är du hänvisad till att själv göra en värdering av de behandlingsstudier som finns.

Börja med att studera »abstract»

En grundregel såväl i den kritiska granskningen som för vardagsläsning-

Sammanfattat

- För svar på frågor om effekten av en viss behandling är randomiserade och kontrollerade studier att föredra.
- En snabb bedömning av tillförlitlighet och relevans avgör om artikeln är värd att läsa.
- En studie har god tillförlitlighet när randomiseringen är adekvat, blindning är nyttjad så långt möjligt och bortfallet av patienter är litet.
- En studie har hög relevans om försökspersonerna påminner om dina patienter och befinner sig i samma vårdmiljö som din egen.
- Förbättringar i patientrelaterade effektmått är viktigare än surrogat-effektmått.
- Slösa inte bort tid genom att läsa studier som saknar tillförlitlighet och relevans.

en är att aldrig läsa artikelns brödtext (introduktion, resultat, diskussion) förrän du etablerat huruvida artikeln är tillförlitlig och relevant för dig och dina patienter.

Alltför vanligt är att läsaren förfors av en lockande titel, hoppar till sista raden i sammanfattningen, »abstract», där finner en spännande slutsats och därefter läser introduktion, resultat och diskussion. Själva metoddelen blir nog ofta styvmoderligt behandlad. Men varför slösa tid på att läsa artikeln om studien har använt en design som är dåligt lämpad för frågan och som hotar att leda till stora risker för systematiska fel, bias? Det är viktigt att tidigt fastställa om studien är dålig och metoden är fel, och att man, oavsett hur spännande resultatet än är, inte kan lita på slutsatserna! Frågan om tillförlitlighet i studiede-

sign och genomförande hänför sig till det som kliniska epidemiologer kallar intern validitet.

Studiens relevans berör aspekter på hur behandlingen givits, hur kontrollgruppen behandlats, på vilket sätt effekten uppmätts och, inte minst, i vilken sjukvårdsmiljö och på vilka patientgrupper studien genomförts. Dvs hur mycket kan jag överföra dessa resultat till mina patienter och min vårdmiljö? Generaliserbarheten av fynden i studien inryms i begreppet extern validitet.

Utifrån de »abstracts» du finner vid din litteratursökning måste du ta ställning till om du skall beställa artikeln i fulltext. Börja med sammanfattningens rubrik och inledning och bestäm dig för om sjukdomen och den specifika behandlingen är av intresse. Gå direkt på metoddelen och bekräfta att studien är randomiserad. Strunta tills vidare i resultat och slutsats. Om studien inte tycks omfatta sådana patienter och sådan behandling som du är intresserad av, eller inte beskrivs som randomiserad, sluta läsa och lägg tills vidare studien åt sidan.

Om, när din genomgång är klar, du inte har hittat några randomiserade studier eller några artiklar som berör exakt den behandling eller patientgrupp du var ute efter kan du tvingas gå tillbaka och söka det näst bästa sättet att besvara din fråga. Grundtanken är att identifiera och använda »bästa tillgängliga vetenskapliga underlag».

Finner du utifrån sammanfattningen att studien tycks vara både tillförlitlig och relevant kan du gå vidare med att beställa artikeln från ditt bibliotek, eller direkt läsa den i fulltext om du har tidsskriften eller dess hemsida tillgänglig.

Använd en granskningsmall

Därefter görs en mer utförlig genomgång av den artikel som tittat sig lovande. I det fortsatta arbetet kan en granskningsmall vara till hjälp. En beprövad sådan redovisas i Figur 1. De två artiklar i serien »Users' guides to the medical literature» som ligger till grund för mallen ger en utmärkt introduktion till kritisk granskning [6, 7]. Mallen kan laddas ner från SBU:s hemsida, där du även kan finna exempel på mallar för diagnostiska studier och systematiska översikter (www.sbu.se). I bokform finns flera utmärkta introduktioner till ämnet; hitintills oöverträffad är Trisha Greenhalghs bok, nu i sin andra upplaga och snart översatt till svenska [8].

Hoppa över artikelns inledande text förutom den sista meningen, som bör läsas noggrant: vad var egentligen studiens målsättning och hur avsåg man att mäta effekten av behandlingen? Är avsikten distinkt formulerad och överensstämmande med din egen fråga? Kom

I. Är resultaten från denna studie tillförlitliga?

1. Randomiserades patienterna till de olika behandlingsalternativen?
2. Var grupperna likvärdiga vid studiens start?
3. Följdes alla patienter upp som randomiserades?
4. Analyserades patienterna i den grupp som de ursprungligen randomiserades till?
5. Var patienter och forskningspersonal blindade avseende behandling?
6. Behandlades grupperna likvärdigt ur alla andra aspekter än interventionen?

II. Vilka resultat visade studien?

1. Hur stor var behandlingseffekten?
2. Hur precis var skattningen av behandlingseffekten?

III. Bidrar resultaten till att ge mina patienter bättre sjukvård?

1. Kan resultaten appliceras på mina patienter?
2. Redovisar studien kliniskt relevanta effektmått, dvs patientrelaterade sådana?
3. Är den sannolika nyttan större än den potentiella skadan och kostnaderna?

Figur 1. Ovanstående frågor ingår i en mall för kritisk granskning av interventionsartiklar (baserad på referens [7]).

ihåg målsättningen till slutet av genomgången: var det verkligen den frågan man senare besvarade? Alltför ofta är utfallet negativt avseende studiens huvudmålsättning (primära slutmåtpunkter, »end-points», eller effektmått) och författarna lyfter då i stället fram något tveksamt fynd avseende annat effektmått eller från en subgruppsanalys.

Fokusera på randomiseringen

Gå nu vidare till metoddelen och studera hur randomiseringen är beskriven. En adekvat randomisering är sannolikt den enskilt viktigaste faktorn för en studies tillförlitlighet [3, 4, 9, 10]. Processen innehåller två moment: att gene-

rera en slumpmässig sekvens som styr i vilken grupp en patient hamnar, och att förhindra att den kliniske provaren får möjlighet att förutse vilken grupp detta blir [11]. Den sistnämnda processen, »concealed allocation», kan säkras genom att provaren först bestämmer sig för att en patient skall inkluderas i studien och därefter ringer upp en central enhet som förvaltar randomiseringslistan, genererad via slumpantal. Man kan också använda garanterat ogenomskinliga kuvert, som öppnas först efter beslut om randomisering.

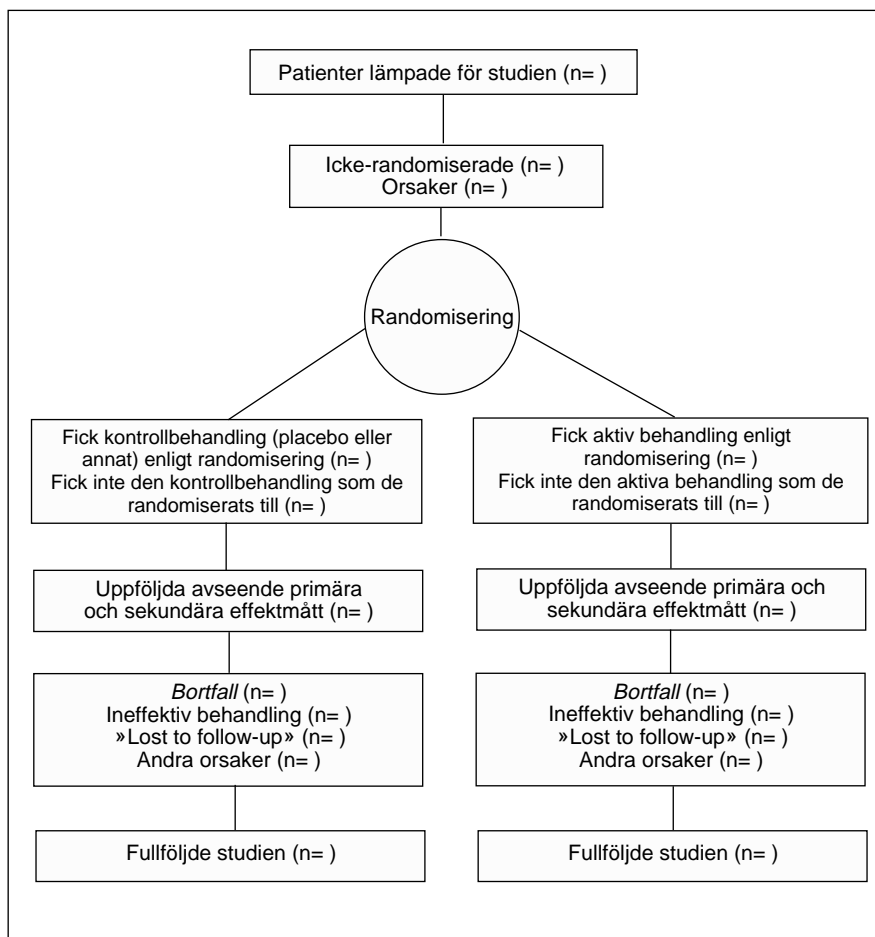
Dessa kan tyckas vara onödiga försiktighetsåtgärder, men det finns empiriskt stöd för och vittnesmål om att enklare förfaranden ofta lett till att studien manipulerats. Studier med oklart randomiseringsförfarande har generellt visat sig ge överskattning av behandlingseffekt jämfört med de studier som använt och redovisat adekvata rutiner för fördelning av patienter mellan aktiv behandling och kontrollgrupp [9]. Dold allokering har inget med blindning att göra utan har som syfte enbart att minimera bias till följd av selektion och störande, »confounding», faktorer. En explicit beskrivning av hur randomiseringen gjorts är en kvalitetsmarkör, medan avsaknad av sådan beskrivning lämnar läsaren i ovisshet. Förfarandet kan ha varit adekvat, men vi kan inte vara säkra.

Baslinjedata centrala för relevansbedömning

Många kliniska studier har undersökt stora mängder patienter för att sålla fram en mindre andel som uppfyllt kraven för att inkluderas i studien. Moderna och välgjorda studier redovisar ofta hur stor andel av patienterna med den aktuella sjukdomen som till slut blev aktuella för randomisering. Det är uppenbart att om man screenat t ex 50 000 hypertoniker för att till slut inkludera 1 000 patienter i studien innebär det att många av våra »vanliga» patienter inte kommit med. Därför bör du ta ställning till om undersökningens exklusionskriterier inneburit att många patienter i en kategori som du själv behandlar har uteslutits.

På vägen mellan inklusion och randomisering är en observationsperiod vanlig för att verifiera att värdet för kolesterol, blodglukos, kroppsmasseindex eller blodtryck faktiskt ligger kvar på den förhöjda nivå som ursprungligen uppmättes. Här faller ofta ytterligare patienter ifrån, eftersom vid en andra provtagning/mätning extrema värden kan ha rört sig mot medelvärdet till följd av »regression towards the mean».

Kliniska karakteristika för de patienter som randomiserats redovisas oftast i en baslinjetabell, stratifierad på grup-



Figur 2. Flödesschema för en randomiserad klinisk studie [13].

pen som fått aktiv behandling respektive gruppen som fått kontrollbehandling. Din största nytta med tabellen är att du kan försäkra dig om att patienterna i ålder, kön, rökvanor eller andra viktiga variabler tycks överensstämma med dina patienter med samma diagnos.

Om så är fallet talar det för att generaliserbarhet till din vardag kan vara god. Men om studien avser hjärtsviktpatienter omkring 60 års ålder, diabetiker med ett genomsnittligt kroppsmasseindex på 38 eller är en KOL-studie med en minoritet rökare bör detta väcka stort tvivel eftersom dessa patienter inte överensstämmer med dem vi möter i vanlig svensk sjukvård.

Det utbredda signifikanstestet för skillnader i baslinjevariabler mellan grupperna är obefogat [10, 12]. Själva slumpmässigheten i en randomisering kan förväntas ge upphov till en del skillnader såväl i gruppstorlek som i baslinjevariabler. Det är meningsfullt att närmare studera endast skillnader i faktorer som har ett självständigt samband med studiens primära effektmått, t ex en studie av trombosprofylax där den aktiva gruppen och kontrollgruppen skiljer sig påtagligt avseende andelen kvinnor som har p-piller- respektive

östrogenerbehandling. En sådan skillnad bör föranleda att man statistiskt korregerar för olikheten mellan grupperna, vilket dock inte kan garantera att felkällan elimineras helt, »residual confounding».

Se upp för bortfallet

En förutsättning för att de systematiska felen skall kunna minimeras är att bortfallet av patienter, från inklusion via randomisering till studiens avslutning, är litet. Vanliga anledningar till att patienter inte fullföljer kan vara rent praktiska och socioekonomiska, speciellt i amerikanska studier med stor rörlighet och där patienterna erbjuds ekonomiska incitament för att delta. Andra orsaker kan vara försämringar i grundsjukdomen (t ex astmatikern som får placebo i stället för steroidinhalationer), varvid bortfallsfrekvensen kan anses vara ett effektmått i sig. Slutligen kan bortfallet vara en följd av besvärande biverkningar av den aktiva behandlingen.

De flesta bortfall är nog inte relaterade till vare sig sjukdom eller behandling och bör alltså fördela sig slumpmässigt mellan grupperna. Om stor andel av patienterna inte fullföljer uppstår dock åter tvivel på studiens tillförlitlighet.

Det är svårt att ange en »varningsgräns» över vilken ett bortfall inte är acceptabelt, men siffran 20 procent figurerar ofta, dock utan empiriskt stöd.

Som läsare bör man kräva att bortfallet är adekvat redovisat separat för de båda grupperna, både i absoluta tal och avseende specifika orsaker. Aktuella rekommendationer om rapportering av RCT föreslår ett flödesschema, i vilket patienternas öden under studien kan följas (Figur 2) [13].

Om man studerar endast de patienter som fullföljde hela undersökningen och helt bortser från dem som efter randomiseringen föll bort baseras resultatet på en speciell patientpopulation, skild från den ursprungliga [14]. Denna studiedesign och analys kallas ofta »efficacy»-studie, eller analys per protokoll. En sådan utforskningsstudie kan vara relevant när man börjar studera hurvida en behandling överhuvudtaget har någon effekt. Generellt överskattar dock en sådan analys den kliniska effekten som ses om behandlingen används på grupper av patienter under mer vardagliga förhållanden.

Om man i stället inkluderar samtliga patienter i den slutliga analysen, oavsett om de fullföljde studien eller ej, får man en så kallad »intention to treat»-analys, eller en pragmatisk studie. Effekten för dem som fallit bort får då värderas från den sista mätning som utförts, eller kan ibland inhämtas från dödsorsaksregister eller andra källor. Om inget är känt om patientens öden kan man förutsätta att han drabbades av det sämsta tänkbara utfallet i de primära effektmåten, en så kallad »worst case»-analys. Det innebär att en analys enligt »intention to treat» alltid kommer att ge en mer konservativ bedömning av behandlingseffektens storlek än en »efficacy»-studie.

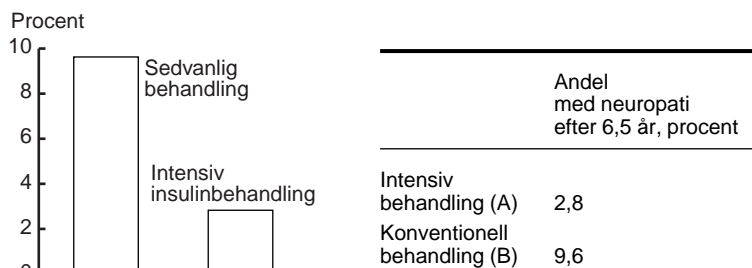
En studie som skall vara användbar för beslut om patienters behandling bör således redovisa resultat enligt »intention to treat», men gärna också som »efficacy». Det är dock i första hand resultatet enligt »intention to treat» som bör ligga till grund för hur studien skall tolkas. Artiklar som inte redovisar, förklarar och tolkar bortfallet har lägre tillförlitlighet.

Blindning

Man hör ofta att »dubbelblindning» krävs för att en studie skall anses vara tillförlitlig. Faktum är dock att kravet på blindning har mindre empiriskt stöd än randomiseringsförfarandet vad gäller en studies interna validitet eller tillförlitlighet. Ändå har jämförelser av studier avseende samma behandling visat att studier utan blindning har ca 17 procent högre effekt än dubbelblinda studier [9].

Begreppet dubbelblind bör utmön-

Neuropatiutveckling



726 unga patienter med typ 1-diabetes utan komplikationer randomiserades till antingen intensiv insulinbehandling eller kontrollgrupp (sedvanlig behandling, dvs 1–2 doser insulin) [18]. Effektmåttet diabetesneuropati uppmätt efter 6,5 år. P-värde för skillnad mellan grupperna <0,001.

Absolut riskreduktion = Andelen med komplikation i kontrollgruppen (B) – andelen med komplikation i behandlingsgruppen (A) = 9,6 – 2,8 = 6,8 procent.

Relativ riskreduktion = Absoluta riskreduktionen/andelen med komplikation i kontrollgruppen = (B–A)/B = (9,6–2,8)/9,6 = 6,8/9,6 = 6,8/9,6 = 71 procent.

»Numbers needed to treat» (NNT) är det antal patienter som måste behandlas under så lång tid som studien pågick för att förebygga ett insjuknande eller ogynnsamt utfall såsom diabetesneuropati, hjärtinfarkt eller död.

$NNT = 1/\text{absoluta riskreduktionen} = 1/(B-A) = 1/(9,6-2,8) = 1/6,8 \text{ procent} = 1/0,068 \approx 15$.

ras. Idealt är det nämligen oftast tre aktörer som bör hållas ovetande om vilken grupp patienten allokerats till: de som utför behandlingen (läkare, sjuksköterska, sjukgymnast m fl), patienten och den som värderar effekten av behandlingen. Den sistnämnda kan vara en läkare som frågar ut och undersöker patienten, en sjukgymnast som värderar rörelseomfång eller smärtupplevelse, en röntgenolog som studerar flebografi eller en patolog som studerar det mikroskopiska preparatet.

I metoddelen bör det klart framgå vilka av dessa tre kategorier som är blindade, och på vilket sätt denna blinding genomförts. Sannolikt är blinding av större betydelse i de fall effektmåttet inrymmer mått av subjektivitet, medan studier som använder död och akut hjärtinfarkt har mindre problem med blinding av utvärderaren. Att hålla patientens grupptillhörighet dold för alla dessa tre kategorier är inte alltid möjligt, men ambitionen bör vara att maximera blindningsproceduren.

Från P-värdenas tyranni till klinisk relevans

Osäkerhet kring tolkningen av statistiska bearbetningar är ofta ett hinder för djupare studier av vetenskapliga redo-

visningar. Inom kritisk granskning och systematiska utvärderingar har däremot analys av de statistiska bearbetningarna ingen central roll. Anledningen är sannolikt att de flesta moderna behandlingsstudier (RCT) har en rättfram och enkel statistisk bearbetning med hypotesprövning av gruppskillnader baserade på t-test eller chi-två-analyser. Problemen uppstår när mer avancerade analyser tillgrips för att visa skillnader i subgrupper, korrigera för olikheter i gruppansammansättning m m. Läkarnas utbildning i statistik har tyvärr fokuserat på signifikansvärden, medan förmågan att tolka resultaten i övrigt sällan tränas. Det är något klinisk epidemiologi vill råda bot på.

Ett enkelt råd är att lägga mest vikt vid studiernas analys av primära effektmått, utan korrigeringar [12]. Strunta i subgruppsanalyser om resultatet inte fallit ut signifikant i hela gruppen. Acceptera inte en »magisk» gräns för signifikans på 0,05 utan att titta närmare på studiens resultat i absoluta termer (se exempel i ruta: Neuropatiutveckling). Hur stor andel i behandlings- respektive kontrollgruppen drabbades av den komplikation man avsåg att förhindra? Skillnaden mellan dessa två tal utgör den absoluta riskreduktionen (ARR),

vilken sällan är lika imponerande som den relativa riskreduktion läkemedelsindustri och entusiastiska forskare väljer att redovisa [6]. »Numbers needed to treat» (NNT) ger det antal som måste behandlas, under så lång tid som studien pågår, för att förhindra att en patient drabbas av komplikationen

Som ett komplement till P-värden bör du leta efter storleken på behandlingseffekten angiven med konfidensintervall, oftast 95-procentiga. Intervallet anger storleksordningen på den effekt som du med 95 procents sannolikhet kan förvänta dig av behandlingen. Dvs om du skulle upprepa samma studie 100 gånger skulle resultatet i 95 av studierna ha fallit inom detta intervall. I en studie skulle behandlingseffekten kunna anges som en absolut riskreduktion på 5 procent (95 procents konfidensintervall, 3–7 procent, P-värde <0,01).

Utifrån absoluta risker får du en uppfattning om riskerna för din patient med placebo eller annan kontrollbehandling, respektive vilken absolut risk som kvarstår även med behandling. Kunskap om dessa risker är oftast betydligt viktigare för dig och din patient än relativa risker och P-värden. Du kan ta ställning till om den signifikanta effekten av behandlingen faktiskt också är av klinisk relevans ur ert perspektiv.

Tolkning och värdering

Till slut får man ta ställning till om studien har tillräcklig tillförlitlighet. Är resultatet i ringa grad påverkade av systematiska fel? Är mina patienter tillräckligt lika dem i studien [15]? Har studien verkligen visat effekt på någon variabel som är av relevans för patientens livssituation och sjukdomslidande, eller har den effekt bara på surrogatvariabler [1, 16]?

Nu är det också dags att reflektera över om andra aspekter som biverkningar beskrivits adekvat. Skulle införande av behandlingen få konsekvenser för vårdorganisationen med avseende på provtagningar, vårdtider, återbesök eller utveckling av vårdkedjor? Finns hälsoekonomiska aspekter, och ger studien möjlighet att besvara frågor i relation till dessa?

Inga absoluta svar

Inom evidensbaserad medicin finns inga otvetydiga svar på hur man skall handla i mötet med patienterna [17]. Att kritiskt sammanställa kunskapsunderlaget ger möjlighet att uttala sig med större eller mindre säkerhet om behandlingens förväntade effekt hos grupper av patienter. I min kliniska praxis behövs jag dock integrera den kunskapen med såväl min kliniska erfarenhet som patientens erfarenheter och preferenser.

På ett högre plan kommer hälsoeko-

nomi, prioriteringar och etik liksom andra värderingar att spela in för behandlingsrekommendationer från läkemedelskommittéer och centrala myndigheter. Ett tillförlitligt och relevant kunskapsunderlag skapar en grund för denna viktiga och svåra process, som fortfarande är outforskad och utvecklade.

Referenser

1. Eliasson M. Evidensbaserad medicin »begins at home». Incitament för bättre beslutsunderlag finns i den kliniska vardagen. *Läkartidningen* 2000; 97: 3860-5.
2. Alton V, Eliasson M. Det är lätt att söka medicinsk information på Internet. Så här gör du – steg för steg. *Läkartidningen* 2000; 97: 4426-34.
3. Jadad A. Randomised controlled trials. A user's guide. London: BMJ Books, 1998.
4. Nilsson G, Sörensen S. Leve den randomiserade studien! Det är den enda metod som med tillräcklig kraft kan värdera nyttan av medicinska åtgärder. *Läkartidningen* 2000; 97: 3445-9.
5. Eliasson M. Den systematiska översikten grundval i evidensbaserad medicin. Ett av de viktigaste tillskotten till den kliniska medicinen under senaste årtiondet. *Läkartidningen* 2000; 97: 2726-8.
6. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. B. What were the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994; 271: 59-63.
7. Guyatt GH, Sackett DL, Cook DJ. Users' guides to the medical literature. II. How to use an article about therapy or prevention. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1993; 270: 2598-601.
8. Greenhalgh T. How to read a paper. The basics of evidence based medicine. London: BMJ Publishing Group, 1977.
9. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995; 273: 408-12.
10. Taube A. Om randomiseringens välsignelser. *Läkartidningen* 2000; 97: 4173-4.
11. Schulz KF. Assessing allocation concealment and blinding in randomized controlled trials: why bother? [editorial]. *ACP Journal Club* 2000; 132: A11-2.
12. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000; 355: 1064-9.
13. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996; 276: 637-9.
14. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *BMJ* 1999; 319: 670-4.
15. Dans AL, Dans LF, Guyatt GH, Richardson S. Users' guides to the medical literature: XIV. How to decide on the applicability of clinical trial results to your patient. Evidence-Based Medicine Working Group. *JAMA* 1998; 279: 545-9.
16. Bucher HC, Guyatt GH, Cook DJ, Holbrook A, McAlister FA. Users' guides to the medical literature: XIX. Applying clinical trial results. A. How to use an article measuring

the effect of an intervention on surrogate end points. Evidence-Based Medicine Working Group. *JAMA* 1999; 282: 771-8.

17. Guyatt GH, Haynes RB, Jaeschke RZ, Cook D, Green L, Naylor C et al. Users' guides to the medical literature: XXV. Evidence-based medicine: principles for applying the users' guides to patient care. Evidence-Based Medicine Working Group. *JAMA* 2000; 284: 1290-6.
18. The effect of intensive diabetes therapy on the development and progression of neuropathy. The Diabetes Control and Complications Trial Research Group. *Ann Intern Med* 1995; 122: 561-8.

Summary

How to critically appraise a randomised trial

Mats Eliasson

Läkartidningen 2000; 97: 5897-5902.

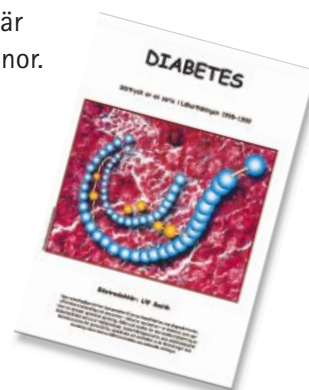
Clinicians need to know how to critically appraise clinical trials in order to choose the best papers upon which to base their clinical decisions. By appraising the section on »methods» the most important questions regarding validity and relevance can be answered. The most important components of validity include adequate randomisation procedures, blinding (where possible) and minimising the number of patients lost to follow-up. The conclusions of a well conducted study should be applicable to a broad range of patients and settings.

Correspondence: Mats Eliasson, Department of Internal Medicine, Sunderby sjukhus, SE-971 80 Luleå.

E-mail: mats.eliasson@nll.se

Nya vetenskapliga rön har lagt grunden till en ny klassifikation, nya diagnoskriterier, effektivare behandling och prevention – inklusive vaccination – av diabetes, som uppvisar en närmast epidemisk spridning. Målen och medlen för den snabba förbättringen av diabetesvården som nu är möjlig belyses i *Läkartidningens* serie, som också analyserar konsekvenserna för patienterna, sjukvården och samhället av de förändringar som Socialstyrelsen hösten 1999 fastställde som nationella riktlinjer.

Priset är 85 kronor.



Diabetes

Beställer härmed.....ex av »Diabetes»

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till LÄKARTIDNINGEN
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Lång väg från evidens till kliniska rekommendationer

|| Evidensbaserad medicin har fått fäste både konceptuellt och i den kliniska verkligheten. Nya och gamla behandlingar ifrågasätts. Gamla därför att effekten har varit oklar, nya därför att de nästan alltid leder till anspråk på ökade resurser. Rapporterna från Statens beredning för medicinsk utvärdering (SBU) har lett till att vissa medicinska åtgärder eller behandlingar som inte befunnits vara tillräckligt väl dokumenterade eller kostnadseffektiva har rangerats ut.

Emellertid har genomslaget för den kunskap som förmedlats i dessa rapporter inte alltid varit så stort som man skulle önska. Blodtrycksbehandling är ett sådant exempel; där har kunskapen om kostnadseffektivaste läkemedelsregim inte nått ut tillräckligt snabbt. Orsakerna till detta har diskuterats; en är naturligtvis marknadsåtgärder från läkemedelsindustrins sida, men det är nog inte hela förklaringen.

Kliniska riktlinjer även för Sverige

Kliniska riktlinjer, »guidelines«, produceras i de flesta länder. Oftast är det de olika professionernas organisationer som står bakom, t ex Royal College of Physicians eller Europeiska kardiologföreningen. Genomslaget för dessa rekommendationer är också nedslående; de behandlingsmål som ställs upp nås sällan till mer än 50 procent. Skälen till detta har också diskuterats.

Riktlinjer utgör sällan lag eller föreskrift som strikt legalt måste följas, även om undantag finns, t ex i Frankrike. Ekonomiska incitament kan också vara kopplade till riktlinjerna, som i de amerikanska HMO-riktlinjerna (Health Maintenance Organizations, de olika försäkringsbolagens organ) eller engelska NICE (National Institute of Excellence, med samband till Ministry of Health och National Health Service). Det sägs att när Margaret Thatcher var premiärminister i Storbritannien blev hon tillfrågad i parlamentet: »What are guidelines? Are they the law that is to be followed?« Hennes svar är typiskt, men beskriver också de flestas syn på riktlinjer: »No they are not the law, but of course they are to be followed!«

Nödvändigheten av att riktlinjer tas fram på ett strukturerat sätt och med god förankring för att åstadkomma legitimitet och följsamhet har diskuterats i Norge (Retningslinjer för retningslinjer) och i European Heart Journal (Legal Implications of Guidelines).

Socialstyrelsen har givits i uppdrag att ta fram kliniska riktlinjer för ett antal sjukdomstillstånd i syfte att bidra till att

Evidensbaserad medicin

utjämna de stora skillnader som finns avseende utfall av och tillgänglighet till vård och behandling, framför allt mellan olika geografiska områden men också mellan olika patientgrupper. Riktlinjerna är i sig inte tvingande, men man skall ha goda skäl att inte använda dem som underlag för de lokala vårdprogram som förväntas bli följden.

Riktlinjerna innehåller avsnitt om epidemiologi, diagnostik, prognos, vård och behandling och består av tre delar med samma kunskapsbas, men med olika utformning beroende på målgrupp. Den mest omfattande delen vänder sig till professionen, de två övriga till patienter/anhöriga respektive administrativa beslutsfattare.

Skall spegla vad som är god vård

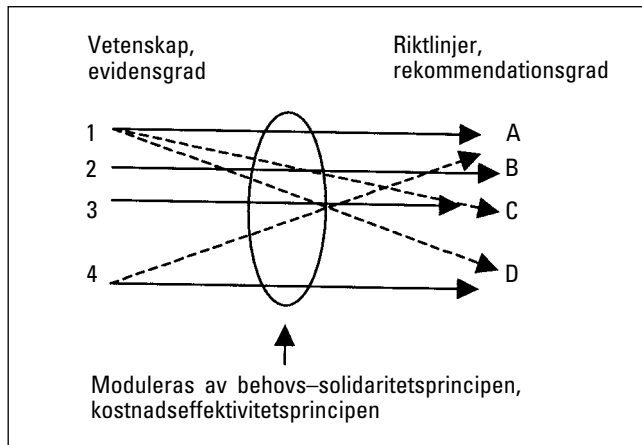
Riktlinjerna är således till för att hälso- och sjukvårdens portalparagrafer skall uppfyllas: God vård på lika villkor; de med störst behov skall ha företräde.

Riktlinjerna skall alltså spegla vad som menas med god vård. En förutsättning för att kunna hävda att det är god vård är att det finns vetenskap eller beprövad erfarenhet som grund. Annan vård får betraktas som experimentell och har sina särskilda regleringar.

Bevis eller evidens för att en åtgärd eller behandling har avsedd medicinsk effekt utgör alltså grunden för att kunna hävda att åtgärden/behandlingen är uttryck för god vård. Det arbete som görs av t ex SBU och Läkemedelsverket för att evidensgradera olika åtgärder/behandlingar är i det sammanhanget mycket värdefullt. SBUs inklusive SBU-Alerts och Läkemedelsverkets rapporter kommer att ligga till grund för Socialstyrelsens fortsatta riktlinjearbete; det första exemplet är underlagen om astma-kroniskt obstruktiv lungsjukdom, som kommer att vara basen för de riktlinjer för dessa tillstånd som nu har börjat tas fram.

Vilka åtgärder skall rekommenderas?

Men bara för att en åtgärd har bevisad effekt, med hög eller medelhög grad av evidens, och därför utgör god vård innebär det inte att den nödvändigtvis kommer att rekommenderas. Den måste också vara angelägen och prioriterad enligt de



Figur 1. Gången från vetenskaplig evidens till kliniska rekommendationer följer inte alltid den heldragna linjen, dvs att stark evidens (1) leder till stark rekommendation (A) och vice versa. Av angelägenhets- eller kostnadsskäl kan en relativt svag evidens utmynna i en stark rekommendation (streckad linje), och, omvänt, en stark evidens i en svag rekommendation.

principer som fastställts av riksdagen i prioriteringsuttalandet. Det vanligaste är nog att hög grad av evidens åtföljs av hög grad av rekommendation, men även åtgärder med låg grad av evidens kan vara rekommenderade om de är mycket angelägna eller kostnadseffektiva. Omvänt kan åtgärder med hög grad av evidens få en låg grad av rekommendation på grund av antingen att åtgärden inte anses så angelägen eller att kostnadseffektiviteten är låg (Figur 1).

Angelägenhetsgraden får väl allmänt anses vara låg för skönhetsoperationer hur väl dokumenterade de än är, och kostnadseffektiviteten är olika för olika läkemedel vid behandling av en rad tillstånd, bl a högt blodtryck, magbesvär, infektioner.

Vem skall avgöra angelägenhetsgraden?

Vem skall då avgöra angelägenhetsgraden? Riksdagens prioriteringsbeslut gjorde ju helt klart att den överordnade principen är den om människovärdet, en tolkning de flesta brukar anse riktig, men börjar man diskutera den kan den ifrågasättas. Ålder är t ex en faktor som i sig inte får användas som prioriteringsgrund, men som ibland används okritiskt. Där emot spelar förstås åldern roll för t ex associerad risk, som då måste vägas mot nyttan (effekten), som för en rad tillstånd är större ju större risken är, dvs ju äldre patienten är. Med åldern följer större risk för komplikationer vid t ex höftoperation, på grund av att patienten också kan ha hjärt-kärlsjukdom. Vidare har förstås ålder betydelse när det gäller återstående förväntad livslängd om en åtgärd vidtas för att rädda livet.

Även andra aspekter av människovärdet, som kön och samhällstillhörighet, kan leda till olikheter i behandling. I riktlinjearbetet måste förstås människovärdesprincipen iaktas, men även könsaspekten bör beaktas om terapieffekten skiljer sig mellan könen.

Behov och solidaritet

Nästa princip är behovs-/solidaritetsprincipen, dvs patienter med mindre behov skall lämna företräde till dem med större medicinska behov. Observera att det här handlar om behov, inte nytta. Här kan riktlinjerna hjälpa till att vägleda framför allt beslutsfattare, så att resurser fördelas till de områden där behandlingen uppvisar god evidens för effekt och där behovet är stort. Det blir här fråga om prioriteringar såväl inom ett område (t ex hjärtsjukdomar) som mellan olika medicinska områden och mellan livshotande och livskvalitetssänkande

tillstånd. Dessa prioriteringar är inte vårdprofessionernas revir; ett medborgar- och befolkningsperspektiv står här i fokus.

Kostnad kontra effektivitet

Den tredje och sista prioriteringsprincipen är den om kostnad kontra effektivitet. Åtgärder av olika karaktär kan ha samma syfte. Rökavvänjningsprogram, t ex, syftar till att rädda liv genom att patienter med kranskärlssjukdom förmås sluta röka. Kranskärlsoperationer kan vara motiverade för att förebygga för tidig död även hos patienter som inte har allvarliga symtom. Det rör sig här om två helt olika program – med samma syfte – som måste vägas mot varandra. Rökavvänjningsprogrammet är mycket mer kostnadseffektivt, men används trots detta inte överallt. Att kranskärlsoperera patienter med instabil angina är också förhållandevis kostnadseffektivt (men ungefär sex gånger dyrare än rökavvänjning). Effekten varierar dock mellan olika grupper av patienter, och införandet av en sådan ny teknologi innebär en stor kostnadsökning. Kompetens och övriga resurser kanske inte heller finns omedelbart tillgängliga.

Hur skall riktlinjerna få genomslag?

Riktlinjerna bör ta hänsyn till dessa faktorer om de skall få ett genomslag hos beslutsfattare och resursfördelare. Än mer komplicerat blir det för beslutsfattare när både angelägenhetsgrad och kostnadseffektivitet för livshotande respektive livskvalitetssänkande tillstånd skall värderas och vägas mot varandra. »Life Years Saved« (YLS), »Quality Adjusted Life Years« (QALY) eller »Disability Adjusted Life Years« (DALY) är användbara mått i det sammanhanget.

Om å andra sidan alla parter är införstådda med och accepterar riktlinjerna har de större chans att bli implementerade utan den fördröjning som annars blir följden och som innebär oacceptabla risker för de högst prioriterade patienterna. Riktlinjerna kan då också utgöra ett av flera underlag för Socialstyrelsens verksamhetstillsyn, dvs tillsyn av en kliniks, ett sjukhus eller landstings program för hur en specifik patientgrupp skall tas om hand, och utfallet för denna patientgrupp. Tillsynen blir då också förutsägbar, eftersom riktlinjerna innehåller rekommendationer avseende såväl indikationsställningar som tillgänglighet och kvalitetsindikatorer.

Riktlinjearbete och evidensgraderingsarbete är resurskrävande och engagerar många människor såväl i de olika professionerna som i patientorganisationer och bland företrädare för finansierare och vårdproducenter. Om detta arbete inte leder till kvalitetsförbättringar i form av bättre behandlingar och bättre rutiner finns förstås ingen anledning att fortsätta. Om evidensgradering inte åtföljs av riktlinjer eller rekommendationer där prioriteringsaspekterna noga beaktats finns det risk för att de prioriteringar som fastställts av riksdagen, och som omfattas av de flesta, blir åsidosatta.

De som bäst behöver den får ingen hjälp, och de åtgärder som har bäst kostnadseffektivitet blir inte använda. •

Povl Riis, professor, dr med, dr med h c, Ældre Forum, København, Danmark (*al-forum@sm.dk*)

Frugtbar opblødning mellem kvantitativ og kvalitativ forskning

Men definitionen af evidensbaseret medicin kunne være bedre forarbejdet

II Man skal ikke ret langt tilbage fra nutiden før medicinens historie dokumenterer, at afstanden mellem sundhedsområdets såkaldt professionelle, dets »medicinmænd«, og befolkningen som helhed var meget kort eller ikke-eksisterende. Med andre ord dominerede, hvad vi i dag vil betegne med folkløriske begreber og beskrivelser. Enkeltiagttagelser af formodede sygdomsårsager eller behandlingsforsøg med naturmidler blev opfattet som bevis for årsagssammenhænge, uden at risikoen for tilfældig samtidighed som forklaring belastede hverken den skriftlige eller mundtlige overlevering, og dermed ej heller den senere praktiske anvendelse.

Indførelse af systematiske observationer kom dermed sent ind i medicinen. Dog indtraf velgennemtænkte tilløb til større klarhed med bl a den engelske marinelæge James Lind i 1747–1753 [1] og danskeren Johannes Fibiger i 1898 [2]. Det var imidlertid først i 1948, at englænderen Austin Bradford Hill og medarbejdere [3], efter et langt teoretisk forarbejde stod for det egentlige gennembrud, at skabe et generelt system for kvantitativ klinisk forskning, i form af den kontrollerede kliniske undersøgelse. Herved skabtes en verdslig vækkelse i form af et opgør med erfaring baseret på enkeltiagttagelser og en tilsvarende dyrkelse af kvantitering og statistisk vurdering af kliniske observationer.

Denne paradigmedannelse inden for sundhedsvidenskaberne var helt nødvendig og har med evolutionære metoder revolutioneret den kliniske forskning verden over, baseret på nye redaktionelle kvalitetskrav, udgivelse af lærebøger [4], postgraduate og nu også prægraduate kurser.

Rammen fik navnet metodologi

Klinisk forskning blev hermed for store områders vedkommende udstyret med et begrebsapparat og en overordnet ramme, der muliggjorde at resultater af en stikprøveundersøgelse med en vis sikkerhed kunne generaliseres til meget store patientgrupper. Rammen fik navnet metodologi (i modsætning til metodelære der har koge bogens natur), defineret som læren om videnskabelige forsøgs planlægning, gennemførelse, fortolkning og publikation [5]. Og selv om termen fastlagdes ud fra det nye kvantitative paradigme spænder det i dag også over de kvalitative metoder i biomedicinen, som dermed på linje med de kvantitative kan og skal kvalitetsvurderes, både på et overordnet og på et konkret, projektorankret niveau.

Evidensbaseret medicin

Se även artikeln på sid 1358 i detta nummer.

Tilsvarende gælder en anden ny term, evidensbaseret medicin (EBM), der fordi den er indført via »sloganeering« har slået an og nu anvendes dagligt, uden at det nødvendige definatoriske forarbejde er gjort. Af mange med en kvantitativ forskningsbaggrund opfattes evidensbaseret medicin således som den kliniske praksis, der hviler på kontrollerede kliniske undersøgelser. Andre, deriblandt undertegnede, inkluderer også resultater af gode kvalitative studier, og endog omfattende erfaringer fra kliniske situationer. Det sidste gælder fx manuelle færdigheder, som ikke etisk eller praktisk kan underkastes det kvantitative paradigme for kontrollerede kliniske undersøgelser, som det især kendes fra dobbeltblinde lægemiddelundersøgelser.

Aksiomer bag alle forsøg på at indsamle troværdig evidens

Forudsætningen for at anvende et udvidet evidensbegreb er imidlertid stor fortrolighed med de aksiomer, der bør ligge bag alle forsøg på at indsamle troværdig evidens i humanbiologien og dens praktiske forvaltning i sundhedsarbejdet.

Det første aksiom udtrykker, at »alt bevæger sig« i en biologisk, tænkende og følende organisme som mennesket, dvs at de enkeltsignaler vi prøver at indkredse i sundhedsfaglige forsøg er omgivet af en meget stor mængde »støj«, stammen fra den mangfoldighed som udgøres af menneskets kompleksitet og biodiversitet.

Det andet aksiom udsiger det næsten fornærmeligt enkle, at alle vurderinger beror på sammenligninger. Dvs at det videnskabeligt er sikrest at skabe et sammenligningsgrundlag via kontrolgrupper, udtaget ved randomisering. Er dette imidlertid ikke muligt i visse forskningssituationer, må et sammenligningsgrundlag skabes på anden måde og den ledsagende bias, forvriddning af tolkningen, beskrives åbent.

Fordelen ved at anvende et overordnet begrebsapparat, der spænder over både kvalitative og kvantitative projekter er også at man mindske den næsten stammekrigslignende tilstand, som har hersket i snart mange år inden for sundhedsfa-

gene og de tilgrænsende discipliner som psykologi, sociologi, antropologi m fl. I et kortvarigt psykologisk perspektiv har tilstanden været let at forklare. Med den kontrollerede kliniske undersøgelses gennembrud så mange af os »lyset« (og ser det stadigvæk), men var også påvirket så stærkt af det nye paradigme, at det blev guldstandard for alle vurderinger af sundhedsfaglige undersøgelser.

Og senere, med en politisk farvet, socialmedicinsk bølge efter at 1968 Nanterre-bevægelsen skyldede hen over fagene, dukkede også kvalitativt baserede undersøgelser op med navn af fx »aktionsforskning«, undersøgelser som i et videnskabeligt univers kun kunne bidrage til en stærk polarisering. Fronten var herefter trukket op. Det var »dem« versus »os« på begge sider.

I de senere år er der imidlertid sket en frugtbar opblødning. En del af de to stammer har lagt våbnene og er begyndt at tale sammen og forstå hinanden i en grad, så samarbejdsprojekter hyppigere og hyppigere optræder i den sundhedsfaglige litteratur. På den ene side erkender forskere skoledet i kvantitative metoder, at disse metoder i et sundhedsfagligt univers ikke kan dække alle problemstillinger, og at også kvantitativt funderede undersøgelser varierer i kvalitet, fra de solidt evidensskabende til de metodologisk invalide; fx at to små grupper ikke udviser nogen signifikant resultatforskel, formuleret af forfatterne som at »behandling A ikke er bedre end B«, selv om testens styrke for den metodologisk kyndige læser er meget lav og type 2-fejlen tilsvarende tårnhøj.

På den anden side, den kvalitative forsknings, oplever man en lignende opblødning som erkendelse af at den kvantitative forskning har sine suveræne forskningsanvendelser, og at spektret af kvalitative metoder og deres praktiske anvendelse også rummer et bredt kvalitetscontinuum.

Ikke kun et nordisk fænomen

Denne udvikling er ikke kun et nordisk fænomen. Den kan registreres i lande som USA og UK, og illustrerer endnu engang at verden i dag ikke er større, end at paradigmeopblødninger eller paradigmeskift optræder multicentrisk i lande på samme generelle udviklingsstrin [6-9]. Det er heller ikke tilfældigt at frontløberne er de almene sundhedsfaglige tidsskrifter, fordi disse ifølge sagens natur belyser mere overordnede problemstillinger, og de meget subspecialiserede de smallere emner, hvoraf de sidste endnu ofte tager udgangspunkt i hvad en forskergruppe metodemæssigt behersker, og i mindre grad hvilke metoder en given problemstilling ville kræve.

Et er imidlertid, at blokerende holdninger er aftagende inden for både kvalitativ og kvantitativ biomedicinsk forskning. Noget andet er, hvad vi kan gøre for aktivt at fremme udviklingen. Afgørende må her være en nyskabt indsats på både præ- og postgraduat niveau.

Prægraduat skal medicinske studenter – og meget gerne andre sundhedsprofessioners uddannelsessøgende, nu da en række fag også er i gang med at opbygge en forskningstradition – introduceres til forskningsmetodologi i bredeste forstand, dvs omfattende både kvantitative og kvalitative metoder. En indføring i alle aspekter inden for indsamling og behandling af både hårde og bløde data vil være uoverkommelig. Men kendskab til grundlæggende principper og de to sæt begrebsapparater bør i nye studieplaner være reglen, så en gensidig sprogforståelse er mulig og så den sproglige opretholdelse af fordomme tilsvarende vanskeliggøres. En radikal nyorientering er tidligere foreslået [10] i form af et scientikum, der dækker det opridsede behov (i øvrigt suppleret af et filosofikum, et psykologikum og et lingvistikum som indledning til alle sundhedsfaglige uddannelser).

Postgraduat er der også et væsentligt pædagogisk behov for at supplere indsigt og begrebsapparat for de mange, der

har deres videnskabelige uddannelse og erfaring på én af de to sider, den kvalitative eller den kvantitative. Atter må det fremhæves at alle ikke skal kunne dække det fulde spektrum, men blot kende begreber og kvalitetskrav nok til at kunne læse kildekritisk henover grænsen, og til at kunne indgå i samarbejdssituationer, hvor et projekts variabler skal hentes på begge sider.

Humanioras betydning for lægeuddannelsen

Det er på denne baggrund meget fortjenstfuldt, at Läkartidningen har taget emnet op, bl a repræsenteret ved Thörn og medarbejderes artikler såvel tidligere [11] som i dette nummer af Läkartidningen, i øvrigt som en parallel til den trend i

Af mange med en kvantitativ forskningsbaggrund opfattes evidensbaseret medicin som den kliniske praksis, der hviler på kontrollerede kliniske undersøgelser. Andre inkluderer også resultater af gode kvalitative studier, og endog omfattende erfaringer fra kliniske situationer.

tiden som fremhæver humanioras betydning for bl a lægeuddannelsen, dvs at ikke blot de kvalitative forskningsmetoder, men at også de hele fag hvis grundlag disse metoder repræsenterer, bliver bragt frem i lyset [12].

Den kontrollerede kliniske undersøgelse har i de sidste 50 år vist sig at være en helt nødvendig betingelse for en pålidelig patientcentreret forskning, men den er ikke altid tilstrækkelig til løsning af alle forskningsopgaver. Her er der inden for et fælles metodologibegreb både behov for, og plads til, også kvalitative metoder.

Referencer

1. Porter R. Ve og vel: medicinens historie fra oldtid til nutid. København: Rosinante, 2000.
2. Gotfredsen E. Medicinens historie. København: Nyt Nordisk Forlag, 1964.
3. Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. BMJ 1948; ii: 769-82.
4. Andersen D, Havsteen B, Riis P, Almind G, Bock G, Hørdér M, eds. Sundhedsvidenskabelig forskning, 5. udg. København: FADLs Forlag, 1999.
5. Riis P. Klinisk og epidemiologisk forskning. Kapitel 2 i reference 4.
6. Green J, Britten N. Qualitative research and evidence based medicine. BMJ 1998; 316: 1230-2.
7. Mays N, Pope C. Assessing quality in qualitative research. BMJ 2000; 320: 50-2.
8. Meyer J. Using qualitative methods in health related action research. BMJ 2000; 320: 178-81.
9. Giacomini MK, Cook D. Users guides to the medical literature. XXIII. Qualitative research in health care: are the results of the study valid? JAMA 2000; 284: 357-62.
10. Riis P. Humanistiske fag og kunst i lægers uddannelse. Ugeskr Læger 2000; 162: 6981-3.
11. Thörn Å. Evidensbaserad medicin med stöd av kvalitativ forskning. Läkartidningen 2000; 97: 3868.
12. Stolt CM. Se hela personen – inte bara patienten. Läkartidningen 2000; 97: 1570-1.

Åke Thörn, överläkare, med dr, yrkesmedicinska enheten, Sunderby sjukhus, Luleå (*ake.thorn@nll.se*)

Marie Bunne, överläkare, med dr, öron-, näs- och halsenheten, Sunderby sjukhus, Luleå (*marie.bunne@nll.se*)

Hans Hallberg, distriktsläkare, med dr, Dalarnas forskningsråd, Falun (*hans.hallberg@dfr.se*)

Kvalitativa forskningsmetoder i klinisk medicinsk praxis

|| Evidensbaserad medicin (EBM) är ett nytt namn på en gammal strävan att i klinisk praktik integrera den enskilde läkarens erfarenheter med bästa aktuella vetenskapliga underlag från externa källor [1, 2]. Begreppet har fått en utbredd förankring och bidragit till ett globalt institutionellt nätverk [3, 4]. Man lyfter där fram den randomiserade kliniska prövningen (RCT) som den metod som har störst vetenskaplig tyngd [1, 5], vilken kan ökas ytterligare genom systematiska översiktsanalyser, om möjligt sammanfattade i gemensamma statistiska mått [1]. Preferensen för en bestämd metod har lett till farhågor att EBM kan föra till ett slutet expertvälde med minskat utrymme för det som kallas läkekonst [6]. Emellertid innefattar begreppet också andra vetenskapliga metoder [7], och även de mest entusiastiska förespråkarna för RCT nämner att »kliniskt omdöme är minst lika viktigt som evidensbaserade vårdprogram« [8].

|| Forskningsmetoder Kvantitativ forskning

Medicinsk vetenskap domineras av kvantitativa forskningsmodeller, vilka har lett till stora framsteg inom såväl grundforskningen som den moderna kliniska praktiken. Mönsterbilden är laboratorieexperimentet, där en studerad händelse/effekt skapas av experimentet samtidigt som alla oberoende variabler kontrolleras och manipuleras av forskaren. Denna strikta form med ett begränsat antal kontrollerade variabler strävar efter att uppnå objektiv information om material, händelser, proportioner och samband.

I den kliniska medicinens komplexa verklighet är det inte möjligt att åstadkomma dessa kontrollerade omständigheter. Kvantitativa modeller för klinisk vetenskap strävar emellertid efter att efterlikna dem genom att skapa så väl reglerade och definierade sammanhang som möjligt. Inom dessa studerar man oberoende variabelers samband med något bestämt utfall på sådana sätt att effekterna kan uttryckas numeriskt [9]. Genom sitt regelverk [8] närmar sig RCT detta laboratorieexperimentella ideal och lyckas ofta framgångsrikt belysa frågor som renodlat rör terapi [2]. Patienter har dock sällan avgränsade problem, som kan omvandlas till specifika frågor och interventioner [10]. Kliniska problem är komplexa och involverar de mest skilda och för forskaren okontrollerbara faktorer; från biomedicinska fenomen till attityder, värderingar och regler [9, 11].

SAMMANFATTAT

Evidensbaserad medicin (EBM) är ett begrepp som betecknar läkarens strävan att i sin kliniska praktik integrera sin erfarenhet med varje given tids vetenskapliga kunskap.

Kliniska problem är i varierande grad sammansatta av de mest skilda faktorer; från biomedicinska fenomen till värderingar, regler och trosuppfattningar. Kvalitativa forskningsmetoder gör det möjligt att systematiskt undersöka och belysa sådana problem och är därför bra komplement till randomiserade kliniska prövningar och annan kvantitativ metodologi.

Frågeställningarna måste vara bestämmande för val av forskningsmetod, och inte tvärtom. Nästan alltid behöver kliniska frågeställningar studeras med utgångspunkt i olika perspektiv.

RCT har fått en dominerande position inom konceptet EBM. Detta kan få olyckliga konsekvenser, varför det tydligare än hittills bör framhållas att RCT blott är ett bland flera möjliga metodval, som syftar till att tillsammans belysa vardagskliniskens problem med bästa möjliga vetenskapliga underlag.

Evidensbaserad medicin

Se även medicinsk kommentar i detta nummer.

Därför kan RCT bara användas för att besvara en mindre del av vardagens kliniska frågeställningar [8].

Kvalitativ forskning

Det kliniska beslutsfattandet kring patienten är en syntes: experterna vetenskapliga bevis tolkas genom den enskilde läkarens

Tabell I. Särdrag inom kvalitativ och kvantitativ forskningsmetodik. (Ur Bunne M. Qualitative research methods in otorhinolaryngology. *Int J Pediatr Otorhinolaryngol* 1999; 51: 1-10, och har tidigare publicerats i modifierad form i *Dagens Medicin* nr 45, 2000).

	Kvantitativa metoder	Kvalitativa metoder
Filosofisk grund	Positivism ¹ , kritisk rationalism ²	Fenomenologi ³ , hermeneutik ⁴
Forskningstradition	Naturvetenskap, till exempel biomedicin, fysik, kemi	Humanism, till exempel antropologi, etnologi, sociologi Omvårdnadsforskning och allmänmedicin
Syfte	Beskriva antal, fördelning eller samband Pröva hypoteser	Identifiera, beskriva och förstå mänskliga företeelser och fenomen. Bilda teorier
Typ av frågeställning	Hur många? Hur ofta? Sammansättning? Orsak och verkan? Frågeställningen är väl avgränsad från sitt sammenhang och studeras till exempel i en experimentell situation eller i epidemiologiska undersökningar	Vad är ... ? På vilket sätt? Karaktär? Innebörd? Frågeställningen studeras som en del av ett större sammenhang, i regel genom fältstudie
Urval	Slumpmässigt urval, randomisering	Strategiskt urval
Datainsamling	Objektiva observationer, mätningar	Intervjuer, observationer
Typ av data	Numeriska	Text. Även bilder och ljud
Analys	Statistiska metoder. Stegvis process	Systematisk kodning, tematisering, sammanfattning och eventuell tolkning genomförs i en cyklisk process
Presentation	Statistik, förklarande text och figurer	Beskrivande text, figurer och tabeller

¹ Positivism utgår från en materialistisk verklighetssyn, som menar att kunskap baseras på våra sinneseffarenheter, »det positivt givna«.

² Kritisk rationalism menar att en vetenskaplig teori aldrig kan bevisas men kan provas empiriskt genom lämpliga experiment och observationer.

³ Fenomenologi försöker att förutsättningslöst förstå meningen och innebörden av mänskliga fenomen och beskriva dem precis som de upplevs, utan anspråk på förklaringar.

⁴ Hermeneutik är ursprungligen läran om tolkning av bibeltexter. Inom kvalitativ forskning syftar hermeneutik på läran om de olika tolkningar som alltid sker vid informationsöverföring mellan människor, via till exempel tal eller text. En kombinerad hermeneutisk-fenomenologisk ansats är vanlig inom kvalitativ forskning.

erfarenhet, läkekonst, intuition, lyssnande och inlevelse [1, 12]. Kvalitativa forskningsmetoder kan användas för att systematiskt undersöka problem som inte är tillgängliga för experimentella metoder [13] och som inte kan frigöras från sina vardagliga, naturliga sammanhang [14], så som ofta är fallet med kliniska vardagsfrågeställningar [9]. Dessa metoder kan därmed bidra till en förbättrad kunskap och förståelse för många förhållanden av betydelse för sjukdomsförloppet som till exempel värk och andra symtom, upplevelser av funktionell kapacitet, sjukdomens inverkan på familj och livssituation [15, 16]. Kvalitativa metoder gör det också möjligt att undersöka attityder, uppfattningar, samspel och processer och kan därför bidra till förståelse för hur evidens omsätts till klinisk praktik [9, 15, 17-19].

Val av forskningsmetod

Benämningen »kvalitativ forskning« är olycklig, därför att man kan tro att denna forskningsgren är av bättre kvalitet än den kvantitativa. Den är dessutom missvisande eftersom det också i den kvantitativa forskningen ingår talrika kvalitativa observationer som kvantifieras. Vid val av studiemetod bör den kliniska frågeställningen vara suveränt avgörande, och därvid finns inga motsatsförhållanden mellan kvalitativ och kvantitativ forskning [20]. Gäller frågan att bedöma effekten av en behandlingsmetod är i regel en kvantitativ metod som RCT att föredra [1, 2, 8, 13, 15, 19]. Gäller frågan att få insikt och förståelse för patienters avvägningar och beslut att följa, eller inte följa, en föreslagen behandling är en kvalitativ metod lämpligare [13, 18, 19]. De två metodområdena bör ses som viktiga och komplementära medel för att belysa och utveckla olika sidor i den kliniska praktiken. (Se Tabell I.)

II Kvalitativ forskning med medicinsk tillämpning

Ofta anges följande skäl till att använda kvalitativa metoder inom medicinsk forskning: för att kunna identifiera, beskriva och tolka symtomupplevelser och förlopp, utveckla begrepp, förstå varför interventioner fungerar eller inte, förklara ovän-

tade resultat i kvantitativa studier och generera hypoteser [21, 22].

Symtomupplevelse och hantering av symtom

För läkaren är symtom ledtrådar som tillsammans med undersökningsfynd används för att kunna ställa diagnos och behandla patienten. För patienten, som upplever symtomen, tolkar dem och relaterar dem utifrån sitt eget perspektiv, har symtomen en annan innebörd. Nedan visas hur en kvalitativ metod används för att kartlägga patienters symtomupplevelser.

Symtomupplevelse. En del personer med kroniskt indragen trumhinna, s k retraktionssjukdom, kan drabbas av plötsliga förvrängningar av ljudet (icke att förväxla med tinnitus). Dessa upplevelser är ibland svåra att verbalisera och går inte att objektifiera med tillgängliga hörselmätningstekniker [23]. Detta kan ha bidragit till att de sällan beskrivs i medicinsk litteratur och således är okända för de flesta läkare. 53 patienter med retraktionssjukdom som haft sådana upplevelser intervjuades med semistrukturerade intervjuer, dvs med öppna icke ledande frågor som kompletterades med fördjupande frågor inriktade på vissa områden, t ex hur ofta, framkallande/lindrande åtgärder, etc [23]. Intervjuerna skrevs ut ordagrant och analyserades med en kvalitativ metod enligt följande: För att få en helhetsbild genomlästes alla intervjuer en första gång. Därefter kodades varje informativt textfragment efter arten av information. Kodad text som speglade liknande information sammanfördes under ett antal teman. Därefter studerades varje tema för sig genom att alla textfragment inom ett och samma tema analyserades. Under denna process framträdde ett antal återkommande mönster som sammanfattades. Slutligen prövades överensstämmelsen mellan denna sammanfattning och innehållet i varje enskild intervju. Undersökningsfynden tolkades dels utifrån patientens perspektiv, dels utifrån ett medicinskt perspektiv.

Två mönster av upplevelser och förhållningssätt. Vanligast var en mycket obehaglig, skärande stark, förvrängd och störande

ljudkvalitet, som personen ofta kuperade genom att medvetet eller omedvetet sniffa in genom näsan. Sådana besvär kunde medföra svårigheter att föra ett normalt samtal, osäkerhet och social isolering samt oro för allvarlig sjukdom, t ex hjärntumör. Oron blev inte mindre av att symtomen ofta missförstods, negligerades eller inte bekräftades av doktorn. Det näst vanligaste mönstret var ett tillfälligt försvagat och dovare ljud, vilket ofta kunde kuperas genom att personen blåste upp luft i örat.

Symtombemästrande förhållningssätt som patogenetiska faktorer. Man brukar förvänta sig konstant försämrad hörsel hos denna patientgrupp. Överraskande nog verkade alltför starka ljud upplevas mest besvärande. Vid sniffning kan ett undertryck skapas i mellanörat, vilket kan leda till kroniska besvär med djupt indragen trumhinna [24]. Individerna kan på så sätt oavsiktligt bidra till sjukdomsutveckling. Fynden motiverar ett nytt och mera differentierat synsätt på utredning och behandling av tillstånd med indragen trumhinna, alltifrån barn med vätska bakom indragen trumhinna till vuxna med retraktionssjukdom. Det är viktigt att noggrant lyssna på patientens symtombeskrivning, bekräfta den och försöka förebygga utveckling av retraktionssjukdom vid sniffbeteende!

Symtombaserade tillstånd

Sedan länge har läkare mött många patienter med medicinskt oförklarliga somatiska symtom och som redan vid första kontakten har en egen diagnostisk etikett på sina besvär. Hit hör sjuka hus-syndromet, multipel kemisk överkänslighet, kroniska smärttillstånd, elkänslighet och kroniskt trötthetsyndrom [25-27]. Tillstånden har många likheter, varför de fått samlingsnamn som »symtombaserade tillstånd« och »den moderna tidens sjukdomar« [26, 27].

De enskilda syndrom som ingår i dessa samlingsbegrepp består alla av diffusa, ospecifika somatiska symtom, som potentiellt kan härröra från flera organsystem. De saknar påvisbara kliniska och laboratoriemässiga förändringar och är vanliga bland annars friska människor [25, 26]. Avsaknaden av kliniska, etiologiska särdrag och påvisbara objektiva tecken gör diagnostiken osäker. Att specificera syndromen genom förfinad analys av enskilda symtom har misslyckats eftersom de upplevs olika av olika individer. Tolkningen av symtom påverkas av många faktorer: förväntningar från familj och samhälle, attityder, uppfattningar och värderingar hos patienter, patientorganisationer och läkare, massmedier, lagar, försäkringssystem m m [26, 28].

Det är svårt att under dessa förhållanden studera dessa syndrom med kvantitativa epidemiologiska metoder. Sådana studier begränsas av grava systematiska fel, dels på grund av ospecifika falldefinitioner som inte objektivt kan valideras, dels på grund av att de föreslagna oklara risk- och orsaksfaktorerna inte medger falsifiering av hypoteser och generalisering av forskningsresultaten [26, 28, 29]. Kunskapen från sådana studier blir vag och allmän och ger föga vägledning till praktisk handling. Att syndromen är bristfälligt förstådda, dåligt definierade och inte kan frigöras från sina sammanhang gör dem dock lämpliga att studera med kvalitativa metoder. Studier med den typen av metodologi kan leda till bättre förståelse av dessa tillstånds uppkomst, utveckling och förlopp [30].

Medicinska behandlingars effektivitet

RCT är sannolikt den bästa metoden för att avgöra medicinska behandlingars verkan. Emellertid räcker inte detta. En påvisad effektiv behandling är verkningslös om inte patienten följer ordinationen.

Kvantitativ forskning. Bristande följsamhet (non-compliance) av läkemedelsordinationer har allvarliga följder för sjukvår-

dens effektivitet och kostnader. Endast ungefär varannan patient med långtidsmedicinering tar sina mediciner i enlighet med läkarens ordinationer [31-33].

Den dominerande forskningsansatsen kring compliance har haft en kvantitativ inriktning och utgått från ett biomedicinskt perspektiv [34]. Man har sökt förklaringar bland annat genom att studera patienters personlighetsegenskaper, medicinska kunskaper om sina sjukdomar, ålder, kön, utbildning, etnicitet och socio-ekonomisk status. Empiriskt har sådana studier endast marginellt förklaringsvärde och har därför inte kunnat bidra till bättre följsamhet av ordinationer.

Kvalitativ forskning. Den kvalitativt inriktade forskningstraditionen försöker förstå hur den som är sjuk upplever sin sjukdomssituation och förhåller sig till de mediciner som ordinerats. Den sjuke betraktas i denna forskning främst som en människa som ingår i ett specifikt sammanhang och med en egen beslutsprocess [35, 36] som grund för att använda eller inte använda ordinerade läkemedel. Utifrån detta synsätt kan patientens bristande följsamhet ses som ett försök att uppnå egen personlig kontroll över sjukdomstillståndet. Detta har studerats i samband med kroniska sjukdomstillstånd som astma [37, 38] och diabetes [39, 40], vilka medför långvarig (livslång) medicinering och återkommande läkarkontroller. Det visar sig att kronisk sjukdom är ett hot mot identiteten [37, 41] med minskat egenvärde och förändrade sociala relationer som följd. Den sjuke ser sin sjukdom snarare som ett existentiellt [42, 43] än som ett medikamentellt problem. En strategi för att avvärja detta hot och bevara den tidigare identiteten som frisk kan vara att undvika den ordinerade regelbundna medicineringen.

En dansk avhandling [44] visar att flertalet av de patienter som besökte allmänläkare ordinerades läkemedel trots att detta inte var önskemålet bakom deras konsultation. I stället var anledningen till konsultationerna oftast en förhoppning om att få en trovärdig förklaring till symtomen. Biomedicinskt (evidensbaserat) »irrationella« motiv bakom läkarens förskrivning var vanligt förekommande. Ibland berodde ordinationen på en ambition att vilja tillmötesgå patientens (av läkaren) förmodade önskan om läkemedel, och stundtals skrevs recept för att snabbt avsluta en tidsödande konsultation i en belastad arbetssituation.

Oro för biverkningar samt rädsla för att bli beroende av läkemedel har i annan kvalitativ forskning visats ha betydelse. I en svensk avhandling [45] angav således var tredje intervjuad patient som långtidsbehandlades på grund av kronisk sjukdom att de betraktade den ordinerade medicinen som »ett gift«. Denna oro gäller ofta läkemedel [37, 46] som läkaren själv uppfattar som riskfria i detta avseende. Det avgörande är emellertid patientens egen bedömning. I denna ingår förväntade positiva respektive negativa läkemedelseffekter som underlag i en subjektivt rationell beslutsprocess [35], vilken kan leda till att ordinerad medicin inte används.

Patientens upplevelse av konsultationens kvalitet har stor betydelse för compliance [47]. En god patient-läkarrelation [48] är en viktig förutsättning för att patienten skall få förtroende för läkaren och följa dennes rekommendationer. Ett förhållningssätt [48, 49] där patienten i en dialog får möjlighet att fritt uttrycka farhågor och förväntningar i samband med konsultationen är här centralt. Detta arbetssätt svarar mot att begreppet compliance (följsamhet) ersätts av begreppet concordance (samstämmighet) [50], som betonar en jämbördig relation där patienten är en partner vars behov av autonomi och aktiva medverkan i beslutsprocessen ses som en självklarhet.

II Kvalitativ forskning i EBM:s metodarsenal

De kvalitativa forskningsmetoderna kan och bör fylla en stor plats inom den kliniska medicinen. De kan bidra till att ut-

forska betydelseaspekter av klinisk praxis som annars inte är vetenskapligt tillgängliga; hur patienter, läkare och anhöriga uppfattar och förhåller sig till symtom, fysiska defekter, smärta, lidande, sjukdomsrisk, behandling, biverkningar av mediciner, processer, kommunikationsmönster och funktion av system [14, 18, 19, 51]. Vi har försökt ge några exempel på hur denna ökade förståelse för sjukdomar, förlopp och effekter av interventioner kan leda till en bättre vård.

Det första exemplet visade hur kvalitativa forskningsmetoder i syfte att beskriva kroppsupplevelser öppnar nya perspektiv för att förstå hur symtom och sjukdom upplevs och hanteras av den drabbade [23, 52]. Förutom att beskriva symptomupplevelsen kunde den kvalitativa ansatsen påvisa beteenden (sniffningsbeteende), som till och med kan bidra till utvecklingen av sjukdomen, i detta fall trumhinneretraktioner. Dessa beteenden hade inte kunnat fångas upp och beskrivas meningsfullt annat än med kvalitativa metoder.

Det andra exemplet visade på den kvalitativa metodiken som ett fruktbart komplement till den epidemiologiska modellens begränsningar vid studiet av de allt vanligare sk »symtombaserade« tillstånden. Förståelsen för dessa svårbegripliga [53] tillstånd kan ökas genom intensiva, deskriptiva studier. Dessa kan leda till klarare grundläggande begrepp och därmed ge bättre förutsättningar för prognostiska ställningstaganden, utveckling av praktiska strategier för prevention, behandling och vidare forskning oavsett metodologisk inriktning.

I det tredje exemplet sätts sjukvårdens effektivitet mycket tydligt i brännpunkten samtidigt som de komplementära dragen i kvalitativa och kvantitativa metoder blir uppenbara. RCT är den idag säkraste metoden för att fastställa en intervention, oftast ett läkemedels, effektivitet. Om inte interventionen accepteras av patienten är dock denna kunskap lite värd. Ofta följs inte ordinationer, och bara i ett litet län som Norrbotten uppskattas oanvänd ordinerad medicin kosta mellan 50 och 100 miljoner kronor per år [54]. Vill vi att den skall användas som ordinerats kan kvalitativa metoder starkt bidra till förståelse för varför resultaten från RCT ofta är så svåra att tillämpa i vardagens kliniska praktik [17].

Tillämpning i klinisk praktik

Lars Werkö anför i sin inledande artikel [12] i Läkartidningens EBM-serie att det är klinikern som skall tolka resultaten av de stora multicenterstudierna och översätta dem till den enskilda patientens situation. Kliniska beslut kring en enskild patient grundas på en syntes av evidens och erfarenhet [55] och av det generella med det individuella [9]. Denna tolkan- de process ingår i vad som ofta kallas intuition, klinisk blick, läkekonst och anses i regel inte vara »forskningsbar«. Detta är en felaktig och onödigt defensiv inställning – kvalitativa vetenskapliga metoder kan och bör användas för att förbättra förståelsen av den komplexa kliniska praktiken [51]. Kraven på deras vetenskapliga grundstruktur skiljer sig inte från de kvantitativa metodernas. I likhet med dessa kräver de grundlighet, konsekvens och sammanhang, neutralitet, öppenhet och kritisk reflektion [18, 51, 56, 57].

Arbeten som presenterar EBM [1, 12] graderar ofta RCT som den främsta vetenskapliga metoden, varefter övriga sorteras in i »en fallande skala av pålitlighet« [5]. Denna rangordning ger intryck av att EBM främst uppmuntrar de kvantifierbara aspekterna på klinisk praktik [58], vilket kan leda till att patient-läkarkontaktens multidimensionella karaktär reduceras till en enkel dikotomi som snarare förvränger än sammanfattar de kliniska prestationerna [10]. Den låter ana en glidning från enbart en rekommendation att använda bästa möjliga vetenskapliga underlag för kliniskt handlande till att också föreskriva vilken vetenskaplig metod som har bästa be-

visvärde [59]. Detta kan få negativa konsekvenser. EBM-seriens redaktör Mats Eliasson [60] redovisar t ex hur ett av seriens referat av en systematisk översikt [61] fått oönskade effekter för delar av diabetesvården i Sverige. Det är mot denna bakgrund vi önskar att de ledande företrädarna för EBM starkt måtte markera att RCT endast är en av många möjliga metoder ur EBM:s arsenal, vari kvalitativa forskningsmetoder borde ingå.

Vilka metoder som skall användas måste styras av frågeställningarna, som i den kliniska praktiken nästan alltid bör undersökas och belysas ur skilda och växlande perspektiv [62]. Stridigheten, som tyvärr finns mellan perspektiven, är konstlad, och den önskar vi bort:

»Jag hörde i drömmen:
– Habib, vill du ha denna lök
eller bara en skiva?
Det föll över mig en stor tveksamhet
Denna gåtas fråga
var mitt livs fråga!
Ville jag ha delen framför det hela
eller det hela före delen?
Nej, jag ville ha båda
såväl delen av det hela som det hela
och att ingen stridighet funnes i detta val.«

(Ur Gunnar Ekelöfs »Diwan till fursten av Engion.«
Stockholm: Albert Bonniers Förlag, 1965.)

Referenser

1. Nordin-Johansson A, Asplund K. Evidensbaserad sjukvård i praktiken. Vetenskapliga grunden bättre än sitt rykte. Läkartidningen 1999; 96:3173-6.
2. Dalén P. Evidensbaserad medicin i vardagsarbetet. Överstepräster som visar oss den rätta, evidensbaserade vägen? Läkartidningen 2000; 97: 497-8.
3. Holman HR. Qualitive inquiry in medical research. J Clin Epidemiol 1993; 46: 29-36.
4. Greenhalgh T. Is my practice evidence-based? BMJ 1996; 313: 957-8.
5. Werkö L. Evidensbaserad medicin. Ett konkret redskap för vårdens verkstadsgolv. Läkartidningen 2000; 97: 2710-3.
6. Feinstein AR. Clinical judgment revisited: The distraction of quantitative models. Ann Intern Med 1994; 120: 799-805.
7. Green J, Britten N. Qualitative research and evidence based medicine. BMJ 1998; 316: 1230-2.
8. Bunne M, Falk B, Hellström S, Magnusson B. The character and consequences of disturbing sound sensations in retraction type middle ear disease. Int J Pediatr Otorhinolaryngol 1999; 51: 11-21.
9. Hyams KC. Developing case definitions for symptom-based conditions: the problem of specificity. Epidemiol Rev 1998; 20: 148-56.
10. Thörn Å. The emergence and preservation of sick building syndrome. Research challenges of a modern age disease [dissertation] Stockholm: Karolinska institutet, Department of Public Health Sciences, Division of Social Medicine, 1999.
11. Nilsson L. Varannan patient följer inte läkarens ordination. Läkartidningen 2000; 97: 4015-9.
12. Hansson Scherman M. Att vägra vara sjuk. En longitudinell studie av förhållningssätt till astma/allergi [dissertation]. Studies in educational sciences 95. Göteborg: Acta Universitatis Gothoburgensis, 1994.
13. Richt B. Mellan två världar. Om konflikten mellan livets krav och doktors önsningar. Lärdomar från familjer med diabetesjuka barn [dissertation]. Studies in arts and science 75. Linköping: Universitetet i Linköping, 1992.
14. Fallsberg M. Reflections on medicines and medication – A qualitative analysis among people on long-term drug regimens [dissertation]. Studies in education dissertations 31. Linköping: Dept of education and psychology, Linköpings universitet, 1991.
15. Malterud K. Behov för nye medisinske forskningsmetoder. Tidsskr Nor Laegeforen 1990; 110: 815-7.

52. Bunne M. Qualitative methods in otorhinolaryngology. *Int J Pediatr Otorhinolaryngol* 1999; 51: 1-10.
53. Malmquist J. Föreställningar om sjukdom. Somatisering-medikaliserings-prioritering. Lund: Studentlitteratur, 2000.
57. Mays N, Pope C. Qualitative research in health care: Assessing quality in qualitative research. *BMJ* 2000; 320: 50-2.
58. Smith BH. Evidence based medicine. Rich sources of evidence are ignored. *BMJ* 1996; 313: 169.
62. Petersson C. Kan man skiva en lök och bevara den hel? *AllmänMedicin* 1999; 20: 184-5.

I Läkartidningens elektroniska arkiv
<http://tarkiv.lakartidningen.se>
 är artikeln kompletterad med fullständig referenslista.

SUMMARY

Qualitative research methods in clinical medicine

Åke Thörn, Marie Bunne, Hans Hallberg

Läkartidningen 2001; 98: 1358-63

Evidence-based medicine (EBM) corresponds to the physician's ambition to integrate his/her experience with the best current scientific knowledge, applying this integrated proficiency to clinical problems. The concept of EBM has become widely accepted and given rise to a global network usually emphasizing the randomized clinical trial (RCT) as the most important scientific method. However, clinical situations are very complex. To a greater or less extent they are all composed of diverse factors, ranging from bio-medical ones to inter-subjective relations, values and beliefs. In addition to RCT and other quantitative procedures, qualitative scientific methods could be used to explore complex clinical problems. Clinical problems always deserve to be addressed from various perspectives. Unfortunately, RCT has acquired such a dominant position as to give the impression that it is the sole preferred method within the EBM framework. Based on three examples, the present article claims that it is the questions asked by an investigation that should determine one's choice of methods, not the reverse. Consequently, we argue that it is important to remember that RCT is just one out of several potential scientific methods. The respective utility of any one method depends on the clinical issue at hand.

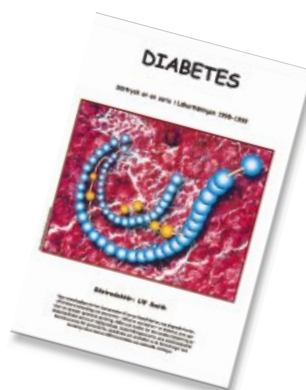
Correspondence: Åke Thörn, Yrkesmedicinska enheten, Sunderby sjukhus, SE-97180 Luleå, Sweden.

Särtryck

Läkartidningen

Nya vetenskapliga rön har lagt grunden till en ny klassifikation, nya diagnoskriterier, effektivare behandling och prevention – inklusive vaccination – av diabetes, som uppvisar en närmast epidemisk spridning. Målen och medlen för den snabba förbättringen av diabetesvården som nu är möjlig belyses i Läkartidningens serie, som också analyserar konsekvenserna för patienterna, sjukvården och samhället av de förändringar som Socialstyrelsen hösten 1999 fastställde som nationella riktlinjer.

Priset är 85 kronor.



Diabetes

Beställer härmed.....ex
 av "Diabetes"

.....
 namn

.....
 adress

.....
 postnummer

.....
 postadress

Insändes till Läkartidningen
 Box 5603
 114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
 under särtryck, böcker

Mätfel i randomiserade kontrollerade studier i neonatalperioden Kvaliteten brister rörande det informerade samtycket

Evidensbaserad medicin

Orvar Finnström
docent, överläkare, barn- och
ungdomsmedicinska kliniken,
Universitetssjukhuset, Linköping
(orvar.finnstrom@lio.se)

II Två nyligen publicerade arbeten som gäller kontrollerade studier i neonatalperioden är av stort intresse för neonatologer men även för andra kliniska forskare. Det första arbetet handlar om kvaliteten i randomiserade, kontrollerade studier, särskilt resultatmått, och det andra arbetet tar upp det informerade samtycket, hur detta ges och tas emot, och om det representerar ett medvetet medgivande.

Mäter vi rätt slutresultat?

I ett arbete av Zhang och Schmidt [1] gjordes en systematisk genomgång av primära resultat(outcome)-mått i randomiserade kliniska studier i nyföddhetsperioden. Författarna påpekar att randomiserade studier är nödvändiga för att man skall kunna göra förutsättningslösa bedömningar av behandlingsresultat. De randomiserade studiernas lov har för övrigt nyligen sjungits i flera artiklar i Läkartidningen. Men är kvaliteten i dessa studier alltid tillräckligt hög så att frågeställningarna (om de finns) kan besvaras?

Enbart randomisering är inte nog

Zhang och Schmidt har försökt belysa denna fråga, och menar att det naturligtvis inte är tillräckligt med bara randomiseringen. Andra betydelsefulla metodologiska spørsmål fordrar också eftertanke, och valet av primärt resultatmått är ett av dessa. Det primära resultatmättet skall vara biologiskt troligt, ha klinisk betydelse och vara potentiellt påverkligt av den intervention som studeras. Vidare har diskreta mått (binära mått, t ex död-levande, synskada-icke synskada) större värde än kontinuerliga mått (t ex mått på ventilation eller syrsättning, Hb, antal blodtransfusioner). Slutligen har tidpunkten, när resultatvariabeln skall mätas, särskilt stor betydelse i neonatala studier.

Vad som kan förefalla vara till nytta tidigt kan visa sig vara katastrofalt senare. Ett bra sådant exempel är det frikostiga bruket av syrgas i neonatalperioden

under 1950-talet som ledde till färre apnéer och ökad överlevnad, men också till en formlig epidemi av det tillstånd som nu kallas för ROP (retinopathy of prematurity) [2].

Författarna har granskat 119 artiklar (av 320 funna), som uppfyllde kriterierna för randomiserade studier, från fem synnerligen välrenommerade tidskrifter (British Medical Journal, Journal of Pediatrics, Lancet, New England Journal of Medicine, Pediatrics). En randomiserad kontrollerad studie definierades som en studie där de nyfödda slumpmässigt valdes ut till en eller flera interventioner. Av dessa 119 arbeten hade endast 91 ett i förväg preciserat primärt resultatmått. Detta var kontinuerligt i 46 fall och diskret i 45 fall. I bara 69 procent av arbetena hade man i förväg definierat minsta möjliga behandlingsskillnad. I den mån man hade preciserat en sådan »minimum detectable treatment difference« så var den orealistiskt stor, i regel 50 procent.

Knappast bättre behandling efter studie

Kan tillämpningen av resultaten från dessa randomiserade studier förbättra behandlingsresultatet för de nyfödda? Författarna ställer sig tveksamma till detta på grund av brister i val av primära resultatmått (man har i stället analyserat en svärm av resultatmått), och på grund av att man ofta använt kontinuerliga i stället för diskreta resultatmått, att studierna med lång uppföljningstid var mycket få och att man haft orealistiska krav på påvisbar behandlingsskillnad. Det sistnämnda innebär att måttliga behandlingseffekter avseende död eller sen morbiditet kan ha missats i många av dessa studier!

Analys av informerat samtycke

Utifrån kunskapen om brister i kvaliteten i randomiserade kontrollerade studier publicerade i mycket välkända tidskrifter är det extra intressant att ta del av erfarenheterna kring det informerade föräldrassamtycket då detta angetts kunna vara ett skydd mot dålig forskning.

Bakgrund. Det är praxis att föräldrar tillfrågas om vetenskapliga studier på sina barn. För att samtycket skall fungera måste föräldrarna vara mentalt kapabla, ha fått adekvat information och ge sitt tillstånd frivilligt. En del tidigare studier antyder att samtycket inte alltid är välgrundat, och det har till och med skildrats som en ritual utan praktisk betydelse.

Ändå krävs detta samtycke i de flesta europeiska länder. Den nu aktuella undersökningen har velat analysera hur processen kring det informerade samtycket fungerar i ett antal europeiska länder, både från föräldrarnas och informatorernas sida.

Metoder. I en välgenomförd multinationell undersökning intervjuades föräldrar till 200 neonatalt sjuka barn angående informerat samtycke. Dessa föräldrar hade tidigare tillfrågats om deras nyfödda barn kunde få delta i en kontrollerad vetenskaplig studie (90 procent av de ursprungligen tillfrågade hade tackat ja).

Efter det att den neonatala studien genomförts intervjuades föräldrarna om hur de upplevt processen kring samtycket. Man använde sig av så kallade semistrukturerade intervjuer, vilket innebär att föräldrar intervjuats utifrån en i förväg bestämd mall, i regel utan fasta svarsalternativ.

Föräldrainsatserna gjordes antingen på sjukhus, hemma eller ibland per telefon, alltid dock av en person som inte kände föräldrarna. De 107 neonatologer, som gjort de ursprungliga förfrågningarna om samtycket, intervjuades också. Undersökningen genomfördes i nio länder (Danmark, Finland, Frankrike, Grekland, Italien, Spanien, Storbritannien, Sverige och Tyskland), varför många bakgrundsfaktorer var olika, men författargruppen ansåg sig ändå kunna dra viktiga slutsatser. Sverige har således varit engagerat i denna studie med 20 föräldrapar och 16 neonatologer. För den svenska delen svarade Nina Nelson och Lena Hellström-Westas.

Analyserna av intervjuerna inriktades på följande kriterier: föräldrakompetens, begripligheten i den information som getts, om föräldrarna förstätt frågorna och om de uppfattat deltagande i studien som frivilligt. Flertalet föräldrar hade fått både skriftlig och muntlig information inför samtycket. Både föräldrar som samtyckt till att låta barnet delta i en studie och sådana som avböjt intervjuades.

Varje intervju analyserades av två oberoende forskare med hjälp av ett strukturerat analysinstrument och ibland av en tredje oberoende person om tolkningsskillnader förelåg.

Resultat. Sju-tio procent av föräldrarna bedömdes ha haft ett eller flera problem enligt något av de fyra analyskriterier

som nämnts ovan. Nästan inga av neonatologerna hade uppfattat samtycket som problemfritt och tveksamheten gällde oftast om föräldrarna förstått informationen.

Fyrtio av de 200 föräldrarna ansågs inte ha varit beslutskompetenta, huvudsakligen beroende på att den känslomässigt svåra situationen med ett sjukt barn, gjort det svårt att fatta beslut. Lika många hade haft svårt att förstå informationen. Huvuddelen av föräldrarna själva tyckte dock att informationen hade varit adekvat. Hälften av dem förlitade sig huvudsakligen på den muntliga informationen, ett fåtal enbart på den skriftliga. Tjugo procent av föräldrarna hade inte varit medvetna om möjligheten att dra sig ur forskningen.

Mycket få föräldrar kände till förekomsten av forskningsetiska kommittéer och deras roll, men en del menade att deras beslutsvända skulle varit mindre om de känt till att det planerade projektet hade granskats av etisk kommitté. Flertalet föräldrar ansåg sig ha tagit beslut om deltagande i studien utan påtryckning, medan över hälften av doktornerna kände tveksamhet på den punkten.

Föräldrar som tillfrågats om medverkan i en akut situation hade naturligt nog mer bekymmer med samtycket än de som tillfrågats i ett lugnare skede.

Föräldrarna uppskattade att forskningen skett efter informerat samtycke, trots att de ibland haft svårt att fatta beslut. De flesta doktorer var också för ett informerat samtycke, men en minoritet ansåg inte att det behövdes och de tillfrågade heller inte föräldrarna.

När det gällde föräldrars skäl att acceptera deltagande i en vetenskaplig studie, så angav de flesta att detta skulle vara av värde för deras barn. Många angav också att det kunde vara av värde för andra barn.

Nästan inga doktorer hade tränats i att informera inför samtycke.

Författarna slutsatser är att det brister en hel del i kvaliteten rörande det informerade samtycket. På åtminstone tre punkter kan en förbättring behövas.

- Läkare behöver tränas i hur man ger information och hur man samtalar angående barnens deltagande i randomiserade kontrollerade studier.
- Forskare som informerar skall inte enbart förlita sig på skriftlig information.
- Föräldrar behöver informeras om de forskningsetiska kommittéernas roll.

Man påpekar att det informerade samtycket är ett viktigt försvar mot deltagande i dålig och samvetslös forskning. Men eftersom det informerade samtycket inte

fungerat så bra måste ansvaret för den här skyddsfunktionen ligga på såväl etiska kommittéer som forskarna själva, anslagsbeviljare och yrkesgrupperna.

De svenska resultaten särredovisas inte, förhoppningsvis sker detta senare.

Kommentar

Även randomiserade kontrollerade neonatala studier, publicerade i tidskrifter med hög status lider således ofta av allvarliga metodologiska brister. Det innebär att viktiga slutsatser av betydelse för barnen ibland inte kan dras, och att potentiellt viktiga fynd döljs vilket naturligtvis inte varit avsikten. Frågan är om dessa brister gäller specifikt för neonatala studier, vilket verkar föga troligt.

Även när det gäller det informerade samtycket så finns det brister. Ett förvånande resultat i Euriconrapporten var att en femtedel av föräldrarna inte uppfattat att man kunde dra sig ur en pågående studie. Anmärkningsvärt var också att mer än hälften av de informerade läkarna kände sig tveksamma till om föräldrarna frivilligt beslutat om deltagande, fastän dessa i regel uppfattat sig ha gjort så. Föräldrar är naturligtvis i en beroendeställning till den behandlande läkaren, vilken sannolikt oftast var den som stod för informationen kring deltagande. Det kan diskuteras om annan än behandlande läkare skall stå för denna information.

I Euriconrapporten [3] lyfter man fram de forskningsetiska kommittéernas roll, och menar att dessa har betydelse för att undvika dålig och för den delen också samvetslös forskning. Zhang och Schmidt [1] visar dock att de forskningsetiska kommittéerna knappast varit särskilt lyckosamma i att förhindra dålig forskning där man använt sig av randomiserade kontrollerade studier.

Frågan är om situationen är annorlunda i Sverige? Jag tror inte heller generellt att de forskningsetiska kommittéerna i vårt land granskar ansökningar tillräckligt noga ur kvalitativ synpunkt, t ex i frågor om primära slutmått och poweranalys, även om en ruta om det senare finns i ansökningsformulären.

En fråga som ställs i rapporten är om det finns situationer där föräldrar inte bör tillfrågas om samtycke. Ett av argumenten har varit att processen kring det informerade samtycket är grymt och svårhanterligt. Resultaten av föräldrainsatserna i Euriconrapporten talar emot ett sådant synsätt. Även om ställningstagandet för dem kan vara svårt, så vill de ändå vara delaktiga. Det andra argumentet har varit att forskningen skulle försvåras eller förhindras. Inte heller för denna ståndpunkt finns det stöd i Euriconrapporten, eller andra arbeten som citeras.

Finns det tillfällen då beslutsprocessen helt skulle kunna läggas, inte på forskarna själva, men på etiska kommittéer, dvs utan att tillfråga föräldrar/anhöriga? Detta diskuteras inte specifikt i Euriconrapporten. Jag är mycket tveksam. Öppenheten och delaktigheten i beslutsförfarandet är oerhört viktigt för ett gott förtroende mellan patient/anhöriga och behandlande läkare.

Referenser

1. Zhang B, Schmidt B. Do we measure the right end points? A systematic review of primary outcomes in recent neonatal randomized clinical trials. *J Pediatr* 2001; 138:76-80.
2. Silverman WA. *Retrolental fibroplasia: a modern parable*. New York: Grune & Stratton, 1980.
3. Mason SA, Allmark PJ for the Euricon Study Group. Obtaining informed consent to neonatal randomized controlled trials: interviews with parents and clinicians in the Euricon study. *Lancet* 2000; 356: 2045-51. Kommentar i: *Lancet* 2000; 356: 2026-7.

Lars Werkö, professor, Statens beredning för medicinsk utvärdering, Stockholm

NICE inte så »nice« som namnet antyder

Smygförsök till terapeutisk byråkrati?

II I England har regeringen gjort flera nya försök att förbättra kvaliteten i sjukvården inom National Health Service (NHS). Ett av dessa är det institut som fått i uppdrag att sammanställa aktuell kunskap och vidarebefordra den till dem som arbetar inom vården, National Institute of Clinical Excellence (NICE). NICE leds av ett litet kansli med professor Michael Rawlins som chef samt består av två stora kommittéer: en Appraisal Committee med 45 ledamöter, en Guidelines Advisory Committee med 24 medlemmar [1].

De flesta i den förstnämnda gruppen är läkare eller ekonomer, men ett antal administratörer, en representant för läkemedelsindustrin (från Glaxo) samt två representanter för patientorganisationer ingår också. Även den andra kommittén har en majoritet av läkare och ekonomer och har dessutom två representanter för industrin (från Merck och Mölnlycke) samt två patientrepresentanter.

Smygransonering?

NICE skall befrämja kvaliteten i sjukvården, bl a genom att ge ut riktlinjer eller vårdprogram baserade på bästa tillgängliga evidens men med hänsyn inte bara till medicinska frågor utan också till kostnadseffektivitet. Detta har lett till att många frågar sig om NICE inte egentligen är ett försök att smyga in ransonering i sjukvården under den falska beteckningen »clinical excellence«. När en engelsk ekonom var i Sverige i september 2000 och presenterade de nya tankarna inom den engelska sjukvården för det möte som IHE organiserat ansåg han att institutet borde ha kallats för National Institute for Cost Effectiveness; det hade också blivit NICE, men hade bättre angett vad det egentligen var fråga om.

I en kritisk ledare i BMJ har tidskriftens redaktör framfört liknande synpunkter [2], vilket har lett till en diskussion i dess spalter [3-10]. En del av inläggen diskuterar principen för det nya institutet, andra diskuterar enskilda beslut som skulle påverka sjukvården. Institutets chef, Michael Rawlins, som nyligen besökte Stockholm för att informera om verksamheten, försvarar naturligtvis både idén bakom dess inrättande och det sätt som man arbetar på [3]. Detta kännetecknas av att man utöver den utvärdering som gjorts av grupper utsedda av NHS organisation för teknologivärdering gör en egen granskning. Denna innebär en bearbetning av tidigare utvärderingar, med hänsyn tagen såväl till ekonomiska frågor som till företags- och patientintressen.

Man torde emellertid hittills knappast ha tillräckliga erfarenheter för att kunna bedöma hur svår denna balansgång

kan bli mellan företag, patienter och det vetenskapliga underlaget.

Kritisk publikation

Redan tidigare har också de reformer som den engelska regeringen infört för NHS utsatts för kritik [11]. I en volym som är sammanställd av tre framstående engelska specialister inom vårdvetenskaperna ifrågasätter man både behovet av ytterligare centrala institut och de motiv som kan ligga bakom reformerna. Även dessa specialister ser risken för underförstådda ransoneringar när det gäller nya och kostnadskrävande metoder för diagnostik och behandling.

Till skillnad från många andra kritiker ser specialisterna – som flera av dem uttrycker sig i denna bok – också vågen av evidensbaserad sjukvård som ett hot mot deras i och för sig berättigade önskan om en fri klinisk vetenskap. En verkligt aktiv klinisk forskning måste, enligt dem, kunna införa nya metoder så snart respektive specialitet eller specialist är övertygad om deras värde.

Detta kommer kanske mest fram i det avsnitt av boken som behandlar kardiologin, en specialitet som excellerat i att inte bara införa nya metoder utan också visat att detta lett till en betydande minskning av för tidig död och sjuklighet i hjärt-kärlsjukdomar. John R Hampton skriver att problemet för kardiologerna inte är brist på evidens, utan svårigheterna att tolka vad de många väl genomförda kontrollerade studierna egentligen visar, och hur deras resultat skall tillämpas i den kliniska verkligheten:

»Does NICE have any better view of the 'truth' than cardiologists who are, after all, specialists in their field?«

Han utvecklar resonemanget med flera aktuella exempel och drar slutsatsen:

»Given the difficulty of interpreting the considerable evidence of cardiovascular therapy, guidelines are likely to reflect the lowest common level of agreement ... We do not need another level of therapeutic bureaucracy to produce this sort of document.« [11, sidan 75].

David G Grahame-Smith, klinisk farmakolog, uttrycker farhågor för att ett organ som NICE kan komma att göra det än svårare att genomföra klinisk forskning. När NICE angett sina riktlinjer fordras det än mer av initiativ och hängivenhet för att driva klinisk forskning av den typ som blivit nödvändig vid införande av nya läkemedel. De stora patientkadrar som kan behövas finns i rutinsjukvården, och om de inte längre kan ingå i forskning rörande nya terapier kommer detta att

omöjliggöra engelskt deltagande i de stora multicenterstudier som initieras när nya läkemedel skall få sin plats definierad i rutinen [11].

Det är emellertid inte bara superspecialister som ser NICE och dess följeslagare CHI (Commission for Health Improvement) som ett hot mot medicinens rätta utövning. Även en representant för allmänmedicin, Brian Hurwitz, uttrycker i samma volym sin oro för att de vårdprogram eller riktlinjer som kommer att diktera sjukvården också kommer att försvåra det förtroendefulla samarbetet mellan primärvårdens utövare och patienterna [11].

Exempel på bearbetning av NICE

Ett exempel på hur NICE med sin »appraisal« kan modifiera den teknologiska utvärderingen kan tas från de arbeten som pågår om behandlingen av Alzheimers sjukdom. De nya läkemedlen för behandling av Alzheimer har nyligen utvärderats av en kanadensisk och en brittisk grupp. Den teknologiska utvärderingen i Kanada [12] uttalar sig om de ekonomiska studierna:

»These studies are all based on modeling, are speculative, and are based on short-term data on efficacy, rather than effectiveness data. However, the modeling technique is the only method currently available to evaluate the drug, in the absence of long-term trials. Long-term prospective studies are required to validate the results demonstrated in these models.«

Den teknologiska utvärderingen i England sammanfattar [13]:

»On the basis of current evidence, the implications of the use of donepezil, rivastigmine or galantamine to treat patients with Alzheimer's disease are unclear. The main issue is whether the modest benefits seen in the outcome measures used in the trials would translate into benefits significant to patients.«

NICE rekommenderar på basis av detta underlag [14]:

»The three drugs should be made available in the National Health Service for England and Wales as one component in the management of people with mild or moderate Alzheimer's disease whose mini-mental state examination score is above 12 points. This is subject to a number of conditions including specialist assessment before therapy is initiated, and a further assessment, which will usually be two to four months after the patient has reached the maintenance dose.«

Däriigenom tillgodoser NICE både företagen och de specialister som är intresserade av att ta hand om dessa patienter. NICE har publicerat sina råd i januari 2001, medan den engelska teknologiska utvärderingen publicerats i mars 2001, vilket visar det intima samarbetet mellan de olika organen i England. Man får förutsätta att det manuskript som NICE fått tillgång till överensstämmer med den rapport som publicerats av NHS Research & Development Health Technology Assessment (HTA) Programme i mars 2001.

Är detta någonting för Sverige?

Per Carlsson har ställt frågan huruvida NICE är någonting för Sverige, dock utan att besvara den [15]. Han framhåller att vi har samma bekymmer som Storbritannien, nämligen avsaknad av adekvata utvärderingar som kan vara underlag för prioriteringar. När det gäller läkemedel gör varken Läkemedelsverket eller Riksförsäkringsverket, som förhandlar om priset på läkemedel, sådana utvärderingar, inte heller Socialstyrelsen när man utarbetar riktlinjer för omhändertagande av

Även specialister ser risken för underförstådda ransoneringar när det gäller nya och kostnadskrävande metoder för diagnostik och behandling. Till skillnad från många andra kritiker ser de också vågen av evidensbaserad sjukvård som ett hot mot deras i och för sig berättigade önskan om en fri klinisk vetenskap. En verkligt aktiv klinisk forskning måste kunna införa nya metoder så snart respektive specialitet eller specialist är övertygad om deras värde.

patienter med vissa sjukdomsgrupper. Man utnyttjar därvid en betydligt mindre organisation än den som står NICE till förfogande. Även om Socialstyrelsen nu också försöker ge råd om prioriteringar lär detta knappast vara tillräckligt som stöd för sådana, eftersom man endast prioriterar inom respektive sjukdomsgrupp. Inte heller gör Socialstyrelsen en så genomgripande värdering av det vetenskapliga underlaget som man gör i England.

Genomslaget av rekommendationerna både i England och i Sverige torde nog vara mer avhängigt av professionens förtroende för dem som utarbetar riktlinjerna än för hur det vetenskapliga underlaget behandlats. En nödvändig utvärdering av dessa verksamheter kommer att ge intressanta resultat, som borde ge ledning för framtiden.

Referenser

1. National Institute for Clinical Excellence: Compilation. Summary of guidance issued to the NHS in England and Wales. London: 2000 vol 1: 9-11.
2. Smith R. The failings of NICE. *BMJ* 2000; 321: 1363-4.
3. Rawlins MR. Reply from chairman of NICE. *BMJ* 2001; 322: 489.
4. Lipman T. NICE and evidence based medicine are not really compatible. *BMJ* 2001; 322: 489-90.
5. Hart JT. Rationing is a respectable word for a disgraceful retreat. *BMJ* 2001; 322: 490.
6. Walker R. »Spinning« is not nice. *BMJ* 2000; 321: 490.
7. Powell M. Latest decision on zanamivir will not end postcode prescribing. *BMJ* 2000; 321: 490.
8. Dhaliwal JS. Editor's choice of acronym for new NICE is Punjabi for thief. *BMJ* 2001; 322: 491.
9. Ellis SJ. Doctors treating patients with multiple sclerosis will lose confidence in NICE. *BMJ* 2001; 322: 491.
10. Robinson S, Valabhji J, Schachter M. NICE's guidance suggests using rosiglitazone in type 2 diabetes later than is ideal. *BMJ* 2001; 322: 491.
11. Miles A, Hampton JR, Hurwitz B, eds. NICE, CHI and the NHS reforms enabling excellence or imposing control? London: Aesculapius Medical Press, 2000.
12. Shukla VK, Otten N, Coyle D. Drug treatments for Alzheimer disease III. A review of pharmacoeconomic evaluations. Ottawa: Canadian Coordinating Office for Health Technology Assessment 2000; technical report 11.
13. Clegg A, Bryant J, Nicholson T, McIntyre L, De Broe S, Gerard K, et al. Clinical and cost-effectiveness of donepezil, rivastigmine and galantamine for Alzheimer's disease: a rapid and systematic review. *Health Technology Assessment* 2001; 5 (1).
14. NICE. *Technology Appraisal Guidance* no 19, Jan 2001.
15. Carlsson P. Är NICE någonting för Sverige? *CMT-Newsletter* 1999;3:2.

Mats Eliasson, docent, Sunderby sjukhus, Luleå (*mats.eliasson@nll.se*)

David Bergqvist, professor, Akademiska sjukhuset, Uppsala

Forskningsresultat bör vara allmänt tillgängliga!

– Fallbeskrivning visar hinder vid kontakt med läkemedelsindustrin

II För att genomföra en systematisk översikt och en värdering avseende effekter och risker med medicinska metoder är tillgång till alla genomförda kontrollerade studier av stor vikt [1, 2]. Långt ifrån alla kliniska studier publiceras och indexeras i bibliografiska databaser, som i t ex Medline. Denna publikationsbias kan yttra sig på så sätt att studierna aldrig publiceras i någon form, enbart redovisas som konferensabstrakt eller i tidskrifter på annat språk än engelska.

Studier som inte visar att behandlingen har effekt (negativa studier) är drabbade av en påtaglig fördröjning innan resultatet publiceras jämfört med positiva studier, om de över huvud taget publiceras [3]. Ansvaret kan vara den enskilde forskarens, anslagsgivande läkemedelsbolags eller medicinska redaktörens. Att inte offentliggöra negativa studier leder till att systematiska översikter överskattar behandlingseffekterna.

Samma problem kan uppstå om vetenskapliga studier som visat effekt publiceras i flera versioner utan att det anges att resultatet redovisats tidigare, s k duplikatpublicering [4-6]. Upprepad publicering står i strid med gällande regler för vetenskapligt arbete, och riskerar att leda till överskattningar av effektstorlek när samma studie flera gånger inkluderas i en systematisk översikt eller metaanalys. Att avbryta en pågående studie leder också till undanhållande av kunskap, och därmed publikationsbias. Det oetiska i detta förfarande har nyligen påpekats [7].

Under arbetet med SBU:s projekt »Förebyggande, diagnostik och behandling av venös tromboembolism« avsåg vi att testa om man kan minska publikationsbias genom att vända sig direkt till aktuella läkemedelsföretag.

Metod och resultat

Enligt Fass fanns det 1999 sex företag som saluförde ofraktionerat eller lågmolekylärt heparin.

I november 1999 fick företagen en officiell begäran från projektgruppen att »tillhandahålla opublicerad dokumentation avseende behandling med de antitrombotiska läkemedel som ni har registrerade i Sverige på indikationen behandling av eller profylax mot djup ventrombos eller lungemboli«. Vår avsikt är inte att peka ut enskilda företag varför de i redovisningen nedan förblir anonyma. Tabell I sammanfattar de tids-

SAMMANFATTAT

Slutsatser och rekommendationer om sjukvårdens metoder ska baseras på bästa tillgängliga evidens.

När negativa studier publiceras mindre ofta än positiva ger systematiska kunskapssammanställningar en falskt gynnsam bild av behandlingseffekten.

Under arbetet med en systematisk kunskapssammanställning tillfrågades sex läkemedelsföretag om opublicerade kliniska studier.

Svaren dröjde mellan 4 och 60 veckor, och flera svar kom först efter ett flertal påstötningar.

Inga opublicerade studier av relevans lämnades ut.

Ett företag underlät att lämna ut en stor negativ studie som förblivit opublicerad till följd av konflikter mellan forskare och företag.

Eftersom det är oetiskt att inte offentliggöra forskningsresultat rörande patienter bör industrin tillsammans med forskare avtala att vetenskapliga studier ska bli allmänt tillgängliga.

Evidensbaserad medicin

fördröjningar som förelåg innan vi fick ett bedömbart svar samt de huvudsakliga resultaten. I Faktarutan framgår detaljerna avseende kontakterna med företagen.

Diskussion

Slutsatser om sjukvårdens metoder ska baseras på bästa tillgängliga evidens. Vår fallbeskrivning avser att redovisa någ-

II Fakta 1

Företag A

En påstötning gjordes per telefon efter fyra månader, varvid en kopia på tidigare skickat brev faxades. Vi utlovades en snar kontakt, men eftersom ingen hörde av sig skrev vi efter ytterligare två månader till den medicinska chefen. Efter en vecka utlovades snar kommunikation, men man påpekade att forskningsavdelningen varit tveksam vad gäller sekretess. Nästa dag levererades en mängd dokumentation hem till gruppens ordförande. Alla dokument skulle hanteras konfidentiellt, men »får användas som referenser till ert arbete om det finns en offentlig presentation eller publikation angiven som källa. I de fall som referensen endast är en internrapport eller är en pågående, ej slutrapporterad studie måste kontakt tas med vårt företag om data ska refereras«. Materialet innehöll ingen ny dokumentation, och gav inga upplysningar av värde. Enstaka rapporter som vi fick i manusform var publicerade när vi fick tillgång till dokumentationen.

Företag B

Efter en månad fick vi svar. Man berättade om en pågående studie, och en avslutad men ännu inte publicerad undersökning. Direktkontakt med den ansvarige forskaren, som hade fria händer från företaget att lämna ut studien, gav besked om två studier, varav den senaste publicerades under år 2000. Studierna var relevanta och hade inte fångats upp i vår egen litteratursökning. Projektgruppen kände dessutom till en brittisk profylaxstudie som företaget var ansvarig för, redovisad vid vetenskapligt möte november 2000, som inte omnämndes i deras svar.

Företag C

Vårt brev besvarades efter tre veckor med e-post, man förklarade sig villig till fortsatt diskussion. Vi förklarade för »Senior Key Account Manager« vilken information gruppen önskade, och gjorde en påstötning via e-post efter tre månader, och via meddelande på telefonsvarare efter ytterligare en månad, utan resultat. Två månader senare skrevs ett brev till den medicinska chefen. Allt gensvar uteblev, varför vi efter ett halvår postade en fjärde påstötning. Redan efter en vecka meddelades kort via e-post att

man skulle skriva till moderbolaget. Ett slutligt svar inkom 60 veckor efter första förfrågan. Man hänvisade i detta till de kliniska prövningar av faktor Xa-hämmare som redan hade presenterats offentligt vid ett vetenskapligt möte några månader tidigare, samt att »för preparat X är alla data som är intressanta publicerade både som artiklar eller åtminstone i abstraktform«. Då projektgruppen hade kännedom om en opublicerad studie avseende profylax med preparat X till sängliggande internmedicinska patienter tog vi kontakt med den för studien huvudansvarige forskaren. Han meddelade att resultatet aldrig publicerats i sin helhet »... due to a conflict between the steering committee and the promoter«, dvs företaget C i sitt hemland [JF Bergmann, pers medd, 2000]. Studieresultatet var negativa och finns kort sammanfattade i ett brev till Lancets redaktör [12].

Företag D

Vi fick inget svar på det första brevet, men en påstötning per brev efter sex månader besvarades med att »man sänt en förfrågan till centrala instanser inom företaget«. Två veckor därefter svarade huvudkontoret att »All studies in the prevention and treatment of venous thromboembolism have been published, at least as an abstract«. På begäran fick vi en bibliografi som sträckte sig till september 1999, och enbart inkluderade redan känt och publicerat material.

Företag E

Vårt brev besvarades inte, men en påstötning efter fyra månader ledde till att vi nästa dag fick abstrakt och föreläsningbilder från två aktuella kongresspresentationer avseende företagets preparat.

Företag F

Inget svar på brevet, men efter påstötning fyra månader senare fick vi omedelbart ta del av en företagsintern och konfidentiell expertrapport om deras preparat. Sammanställningen tillförde inte något nytt.

Tabell I. Sammanställning av resultatet av projektgruppens förfrågan till läkemedelsföretagen avseende opublicerade data.

Företag	Tid till slutligt svar (veckor)	Antal påstötningar	Översänt material	För gruppen nya data?
B	4	Ingen	Namn på ansvarig forskare	Ja, två nya, men publicerade, studier
E	17	1	OH-bilder från kongress, abstrakt	Nej
F	17	1	Produktresumé	Nej
A	23	3	Produktresumé	Nej
D	28	1	Bibliografi	Nej
C	60	4	Inget	Nej

ra hinder som finns för att leva upp till denna målsättning. När projektgruppen försökte undanröja en sannolik publikationsbias genom direktkontakt med de sex tillverkarna tog det mellan 4 och 60 veckor att få svar. Alla, utom ett företag, krävde minst en påstötning för att besvara en förfrågan från en svensk sjukvårdsmyndighet, och i extremfallet krävdes fyra påstötningar och dröjde 60 veckor.

Det är viktigt att påpeka att vi inte hävdar att utfallet skulle vara representativt för de drygt 130 företag som marknadsför läkemedel i Sverige. Vi inser också att det i enskilda fall

kan finnas rimliga förklaringar till de svårigheter vi mött. Ändå är resultatet nedslående, och pekar på ett stort behov av gemensamma insatser från ansvariga forskare, anslagsgivare, medicinska tidskrifter och instanser med ansvar för övergripande slutsatser och rekommendationer till svensk sjukvård.

Resurskrävande, utan konkreta resultat

Det material som ställdes till förfogande var av mycket varierande kvalitet och omfattning. Medan företag A lämnade en

stor uppsättning material, företag F en intern och omfattande produktresumé och företag B pekade ut studier under publicering, så bidrog inte övriga med något relevant material. Några skickade oss abstrakt från konferenser, vilka kan vara till viss hjälp genom att peka på vad som eventuellt senare kommer att publiceras.

Inget material förändrade gruppens arbete eller kommande slutsatser. Den viktigaste erfarenheten är att företagen inte bidrog med något helt opublicerat material av värde för gruppens översikt. Processen var tids- och resurskrävande.

Tre iakttagelser är värda ytterligare kommentarer.

Anmärkningsvärd fördröjning

För det första är det mycket anmärkningsvärt att endast ett av sex företag besvarade vår begäran utan påstötning eller tidsfördröjning, och orsakerna till agerandet är oklara.

Möjligen rörde vår förfrågan vid svåra och centrala ämnen i relationen mellan industri och offentlig sjukvård, där dessutom upparbetade rutiner för ett sådant samarbete saknas. En välvillig tolkning är också att de stora strukturella förändringar som läkemedelsindustrin nyligen genomgått har lett till oklarheter avseende ansvar och befogenheter vid de svenska kontoren.

Viktiga kliniska studier är opublicerade

För det andra är det uppenbart att företag C underlåtit att redogöra för en studie som visade att deras preparat inte hade avsedd effekt, en studie av betydelse för gruppens kommande slutsatser. Det är knappast troligt att huvudkontoret saknade kunskap om studien, som utförts i samma land, utan vi tolkar detta som undanhållande av oönskade forskningsresultat.

Vi kan inte veta om liknande studier döljer sig hos andra företag, eftersom det inte finns register över kliniska studier. Nationella studier av marknadsföringskaraktär är sällan kända av företagen centralt, och bör definitivt sammanställas inom företaget och kunna redovisas. Ett omvänt problem är att den nationella grenen av företaget inte alltid har klart för sig vilka data som finns centralt, och det är möjligt att vårt resultat blivit annorlunda om vi direkt tillskrivit huvudkontoren.

Sekretess gynnar inte patienterna

För det tredje krävde två företag konfidentiell hantering av data som lämnades ut. Vi har förståelse för behovet av företagssekretess, men om vi inte skulle få redovisa och tolka materialet är det knappast meningsfullt att få tillgång till det.

Rättsläget tycks här oklart, men man kan konstatera att Läkemedelsverket har tillgång till alla kliniska provningar, publicerade såväl som opublicerade, som genomförts före godkännande av preparatet. Verket kan lämna ut data efter individuell provning.

Nyligen redovisade forskare från Läkemedelsverket en jämförelse mellan de opublicerade studier som företagen lämnat in för registrering av SSRI-preparat och de av dessa studier som sedermera publicerats [8]. De studier som företagen publicerat överskattar behandlingseffekten jämfört med det totala underlaget. För ett icke namngivet preparat var andelen som svarade på behandlingen 10 procentenheter större i de publicerade studierna än en i sammanvägd analys av samtliga inlämnade studier.

Opublished forskning är oetisk

På samma sätt som dålig forskning är oetisk är forskning som inte är offentligt tillgänglig oetisk. I den senaste versionen av Helsingforsdeklarationen hävdas att »The design of all studies should be made publicly available« [9], vilket är ett starkt

stöd för en prospektiv registrering av alla kliniska studier [10]. Än viktigare är skrivningen i paragraf 27: »Negative as well as positive results should be published or otherwise made publicly available«. Självklart kan man inte acceptera att patienter utsätts för kliniska provningar med potentiella risker utan att kunskapen om resultaten bidrar till vetenskaplig och klinisk utveckling.

Avtala om skyldighet att redovisa studier och resultat!

Det danska läkarförbundet slöt nyligen ett avtal med den danska läkemedelsindustrin som uttryckligen lägger fast att båda parter har rätt att publicera resultat av kliniska studier oavsett resultatet [11]. Enligt avtalet ska man sträva efter att även negativa studier ska publiceras. De kliniska forskarna avråds från att delta i studier där detta avtal inte följs. Liknande avtal har slutits i Storbritannien, där åtminstone studieupplägget för alla studier kommer att offentliggöras när ett preparat godkänts på den brittiska marknaden.

I Sverige har Läkemedelsverket nyligen föreslagit att de etiska kommittéerna, inför ett godkännande av en planerad studie, borde kräva att det i studieprotokollet framgår att företaget kommer att ange att studien är genomförd, t ex på sin hemsida. Dessutom bör företagen, enligt detta förslag, göra resultatet allmänt tillgängligt, åtminstone för godkända läkemedel [Björn Beerman, pers medd, 2001].

Vi stöder detta förslag, och vill också uppmana Läkarsällskapet och Läkarförbundet att initiera ett avtal, som i Danmark, så att även svenska forskare har garantier för att resultaten av studier görs allmänt tillgängliga, oavsett resultatet. Samma bör gälla för forskning som finansieras av allmänna medel eller olika stiftelser där den ansvariga forskaren ibland avstår från att publicera sina negativa resultat. Ett öppet register över pågående svenska kliniska studier borde lätt kunna upprättas med basen i etikkommittéernas godkännande.

Vi vill med denna artikel bjuda in berörda parter till diskussion om hur vi gemensamt, forskare, industri och myndigheter, kan samverka för att synliggöra bästa tillgängliga vetenskapliga kunskap, som grund för beslut om sjukvårdens metoder. De parter som deltar i en sådan utveckling har alltså vinna på öppenhet, medan de som ställer sig utanför måste misstänkas ha brister i sin vetenskapliga dokumentation. Helsingforsdeklarationens nya skrivelser banar nu väg också för en internationell samverkan i dessa frågor.

Referenser

1. Mulrow CD. The medical review article: state of the science. *Ann Intern Med* 1987;106:485-8.
2. McAlister FA, Clark HD, van Walraven C, Straus SE, Lawson FM, Moher D, et al. The medical review article revisited: has the science improved? *Ann Intern Med* 1999;131:947-51.
3. Stern JM, Simes RJ. Publication bias: evidence of delayed publication in a cohort study of clinical research projects. *BMJ* 1997;315:640-5.
4. Johansen H, Göttsche P. Problems in the design and reporting of trials of antifungal agents encountered during meta-analysis. *JAMA* 1999;282:1752-9.
5. Rennie D. Fair conduct and reporting of clinical trials. *JAMA* 1999;282:1766-8.
6. Eliasson M. Duplikatpublicering ett sätt att försköna forskningsresultat. Oetiskt missbruk som hotar trovärdiga systematiska översikter och metaanalyser. *Läkartidningen* 2000;97:3454-6.
7. Bergqvist D. Industriebeslut att stoppa studier – viktigt att informera om att och varför så sker. *Läkartidningen* 2000;97:5367-8.
8. Ahlqvist-Rastad J, Melander H, Beerman B. Evidence of publication bias for antidepressant drugs (Poster D29). In: *Cochrane Colloquium*; 2000. Cape Town; The Cochrane Collaboration: 2000.

9. Jakobsson LH. Handen på hjärtat – har du läst Helsingforsdeklarationen. Ny version av den etiska grundbulten i medicinsk forskning. *Läkartidningen* 2001;98:12-4.
10. Tonks A. Registering clinical trials. *BMJ* 1999;319:1565-8.
11. Samarbejdsaftale om kliniske forsøg mellem lægestand og lægemiddelindustri. Den Almindelige Danske Lægeforening, september 2000. URL: <http://www.dadlnet.dk/2laegeforeningen/retogetik/vejledninger/dadlvejledninger/klinisk.htm>
12. Bergmann JF, Caulin C. Heparin prophylaxis in bedridden patients. *Lancet* 1996;348:205-6.

SUMMARY

Results of research should be freely accessible!
– Case study demonstrates obstacles to contact with the pharmaceutical industry

Mats Eliasson, David Bergqvist
Läkartidningen 2001;98:3913-16

Recommendations on the use of medical technologies should be based on best available evidence. Publication bias due to delayed publication or non-publication of negative studies leads to inflated effect sizes in systematic reviews. We carried out a health technology assessment on prevention and treatment of venous thromboembolism, and in order to minimize publication bias asked the six pharmaceutical companies with antithrombotic drugs registered in Sweden to supply us with results from unpublished clinical trials. The answers were returned after 4 to 60 weeks, and with only one exception, after multiple reminders by telephone, fax and e-mail. No relevant unpublished studies were reported to us although we were made aware of two published studies that our literature search had not identified. One company did not mention a large negative unpublished clinical trial with their drug which has remained unpublished due to conflict between the company and the steering committee. We conclude that the companies did not contribute much and that the process was very time consuming with little result.

As non-publication of trial results is unethical, we propose agreements between industry and researchers that all studies be prospectively registered and made publicly available.

Correspondence: Mats Eliasson, Sunderby Sjukhus, SE-971 80 Luleå, Sweden (mats.eliasson@nll.se)

Särtryck

Läkartidningen

När Försäkringsmedicinska Sällskapet bildades för att främja försäkringsmedicinens utveckling samlades 14 artiklar publicerade i *Läkartidningen* 1996 till ett särtryck. Detta belyser hur försäkringsläkare arbetar inom allmän och privat försäkring och tar upp försäkringsmedicinska problem från patientens synvinkel.

Riskbedömning vid barnförsäkringar, etiska problem i samband med gentestning och försäkring, samt de kniviga ärenden som gäller nack-skulderbesvär, inklusive pisksnärtskador, behandlas bland annat i artiklarna.

Priset är 50 kronor.

Försäkringsmedicin



Beställer härmed ex
av "Försäkringsmedicin"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till Läkartidningen
Box 5903
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Enligt den kunskapen kan fissurer i den yttre delen av anulus, diskbråck, läckage av diskvävnad, diskdegeneration och facettledsartros [8] kanske vara orsaken till ryggsmärtan hos majoriteten av personer med kroniska ryggsmärtor.

rats här borde vi kanske spendera mer tid på att förebygga ryggbesvär, lära oss hur vi effektivt skall behandla de strukturer som är den riktiga orsaken till ryggsmärtan, och spendera mindre tid på att massera, manipulera, värma och kyla ner strukturer som har mycket lite att göra med alstringen av ryggsmärtan.

Referenser

1. Kuslich SD, Ulstrom CL, Michael CJ. The tissue origin of low back pain and sciatica: a report of pain response to tissue stimulation during operations on the lumbar spine using local anesthesia. *Orthop Clin North Am* 1991;22:181-7.
2. Bogduk N, Tynan W, Wilson AS. The nerve supply to the human lumbar intervertebral discs. *J Anat* 1981;132:39-56.
3. Yoshizava H, O'Brien JP, Thomas-Smith W, Trumper M. The neuropathology of intervertebral discs removed for low-back pain. *J Pathol* 1980;132:95-104.
4. Sachs BL, Vanharanta H, Spivey MA, Guyer RD, Videman T, Rashbaum RF, et al. Dallas discogram description. A new classification of CT/discography in low-back disorders. *Spine* 1987;12:287-94.
5. Vanharanta H, Sachs BL, Spivey MA, Guyer RD, Hochschuler SH, Rashbaum RF, et al. The relationship of pain provocation to lumbar disc deterioration as seen by CT/discography. *Spine* 1987;12:295-98.
6. Simmons JW, Emery SF, McMillin JN, Landa D, Kimmich SJ. Awake discography. A comparison study with magnetic resonance imaging. *Spine* 1991;16 6 Suppl: 216-21.
7. Olmarker K, Rydevik B: Exponering för diskvävnad sensitiserar nervrötterna. *Läkartidningen* 1998;95:5618-22.
8. SBU. Ont i ryggen – orsaker, diagnostik och behandling. SBU rapport 1991. Stockholm: Statens beredning för medicinsk utvärdering, 1991.

Förbättra presentationen av patientnära forskning!

Den patientnära forskningen har i många år kritiserats för dålig vetenskaplig metod, små randomiserade studier som saknar vetenskaplig styrka samt undermålig presentationsteknik. Den experimentella forskningen kan också kritiseras, där t ex försöksdjur »inkluderas tills man når signifikanta resultat« och publicerar dem då. Den beräkning av styrka (power) som alltid görs i stora randomiserade studier saknas ofta såväl i den patientnära som i den experimentella forskningen.

LARS PÅHLMAN

professor, kirurgiska kliniken, Akademiska sjukhuset, Uppsala
Lars.Pahlman@kirurgi.uu.se

II Med detta som bakgrund vill jag lyfta fram mina erfarenheter som läsare av sammanfattningar (abstracts), dels till Svensk kirurgisk förenings »kirurgvecka« inom ramen för kolerektalkirurgernas förening, dels till ett flertal internationella kongresser.

Metodiken lyser med sin frånvaro

När man på kort tid försöker tränga in ett arbete slås man av avsaknaden av vetenskaplig metodik. Bakgrundsdiskussionen och syftet med studien framgår ofta ganska väl. Metoden är långt ifrån så väl beskriven att det går att förstå hur man har tänkt.

Resultatbeskrivningen är ej heller optimal och inte så sällan ganska medioker där det absoluta »lågvattnmärket« är när resultatet beskrivs i en enda rad: »resultat kommer att presenteras på mötet«.

Att tro att en dylik sammanfattning skall kunna accepteras är naturligtvis helt aningslöst. För att man över huvud taget skall kunna bedöma ett vetenskapligt arbete måste det finnas ett syfte, en metod som är begriplig, ett resultat samt en konklusion som besvarar syftet och målet med studien.

Tyvär är det många som slarvar även i sistnämnda avseende och inte svarar på data man presenterar och drar de rätta konklusionerna. Många gånger görs mycket vidlyftiga konklusioner som ej går att styrka i framställningen.

Saknar genomtänkt strategi

Av sammanfattningarnas kvalitet att döma saknar jag en genomtänkt strategi i många av de kliniska studier vi gör. Det finns alltid undantag där man vill experimentera och visa resultat av en ny metodik, där det kan finnas skäl att göra avkall på

Evidensbaserad medicin

den vetenskapliga ansatsen för att sprida en nyhet. Men dessa undantag är få och majoriteten av de studier vi gör borde ha föregåtts av ordentlig genomgång innan man gör studien, och inte som jag upplever det nu, data samlas in och fynden blir hypotesgenererande. Det måste ju finnas en hypotes från början som man skall besvara.

Våra vetenskapliga sammankomster

Viktigt är att vi har en hög kvalitet på våra vetenskapliga sammankomster. Många föredrag brukar vara bra, men det

Det absoluta »lågvattnmärket« är när resultatet beskrivs i en enda rad: »resultat kommer att presenteras på mötet«. Att tro att en dylik sammanfattning skall kunna accepteras är naturligtvis helt aningslöst.

är väl så viktigt att kunna göra en sammanfattning som visar hur man tänkt och vilka resultat som noterats. Detta är viktigt då den kliniska patientnära forskningen ibland hamnat i »skottgluggen«. För att vi skall kunna få medel till denna viktiga del krävs att rapporteringen görs professionellt. •

(En reviderad version av artikeln har tidigare publicerats i Svensk Kirurgi nr 3/01.)

Marianne Sullivan, professor i psykologi, Göteborgs universitet, chef, sektionen för vårdforskning, institutionen för invärtesmedicin, Sahlgrenska universitetssjukhuset, Göteborg (*marianne.sullivan@medicine.gu.se*)

Alf Tunsäter, leg psykolog, fil dr, specialist i allergologi, överläkare i medicin, Närsjukvården Österlen, Simrishamns sjukhus

Hälsorelaterad livskvalitet informativt effektmått i kliniska studier

Tillämpningsexempel: astma/kroniskt obstruktiv lungsjukdom

■ När behandlingsresultat vid t ex astma/kroniskt obstruktiv lungsjukdom (KOL) ska värderas kan FEV₁ och andra test av lungfunktionen liksom blodgasanalyser ses som surrogatmått för det som spontant upplevs som mest angeläget i patienternas liv: att slippa svår symtombörda, sjukhusvård och tung mediciner; att överleva och att ha en god livskvalitet, dvs att må bra och fungera i vardagen. En nyligen slutförd systematisk kunskapssammanställning av behandling av astma och kroniskt obstruktiv lungsjukdom har tagit fasta på detta [1]. Artikeln är en bearbetning av valda delar av denna SBU-rapport.

Incitamentet till att livskvalitetsforskning utvecklats inom hälso- och sjukvården är »hälsans paradox«: dvs förbättrad hälsoliv enligt traditionella indikatorer åtföljs inte automatiskt av bättre funktionsförmåga i vardagen, välbefinnande och tillfredsställelse med hälsan. Utvärdering av den givna vården kräver givetvis allmänt accepterade indikatorer. Livslängd, dödlighetstal, data från sjuk- och skaderegister och traditionella medicinska effektmått, såsom laboratoriedata, läkarbedömd sjukdomsaktivitet etc, har här sin givna plats. Sådana indikatorer är dock otillräckliga för att förstå hur människors livskvalitet påverkas av ohälsa och vilka vinster och begränsningar medicinska omhändertaganden medför i patienternas ögon.

Livskvalitet: begrepp och mätinstrument

Begreppet livskvalitet är, liksom hälsa, mångtydigt och kan inte ges en distinkt definition. Båda begreppen avspeglar olika aspekter av välbefinnande, men livskvalitet har ett avsevärt bredare innehåll. Hälsorelaterad livskvalitet innebär en pragmatisk avgränsning och avser främst funktion och välbefinnande vid ohälsa, sjukdom och behandling. Modellen för begrepp och mätning kan liknas vid ett kontinuum som löper från sjukdomsspecifika till generella aspekter, från fysisk till psykisk hälsa och från funktionsinskränkning till välbefinnande. Här ges exemplet astma/kroniskt obstruktiv lungsjukdom (Figur 1).

Utvecklingen under 1990-talets senare del har inneburit att standardiserade livskvalitetsmått börjat användas i kliniska studier och läkemedelsprövningar. Härmed kommer kunska-

SAMMANFATTAT

Tillförlitlig metodik för mätning av hälsorelaterad livskvalitet har utvecklats och prövats under ett par decennier i internationellt samarbete mellan läkare och metodexperter. De vanligast förekommande formulärens har anpassats och prövats för svenska förhållanden.

Frågeformulären besvaras av patienten själv och omfattar väldefinierade nyckelområden: symtombörda, funktionsinskränkningar och välbefinnande fysiskt, psykiskt och socialt.

Hälsorelaterad livskvalitet som effektmått kan ge vägledning i dimensioneringen av behandlingsstudier, då denna typ av mått har dokumenterad förmåga att styrka kliniskt meningsfulla förändringar i hur patienten mår och fungerar i vardagen.

Hälsorelaterad livskvalitet lämpar sig också väl som effektmått vid systematiska kunskapsöversikter. Den viktigaste orsaken är att patientrapporterad information har ett mervärde, såsom visas i exemplet astma/kroniskt obstruktiv lungsjukdom (KOL).

Evidensbaserad medicin

pen att öka om den kliniska tolkningen av utfallet i dessa instrument och om det mervärde patientbaserad information kan ha. Ett problem har varit att få, om ens några, studier dimensionerats med särskild hänsyn till potentialen hos livskvalitet som effektmått, en annan att många specifika mått ännu inte validerats internationellt.

De här framförda exemplen på användbara formulär från astma/KOL-området innebär att specifika mått kombineras med generella för god täckning av begreppet hälsorelaterad livskvalitet. Så t ex är SF-36 Hälsoenkät det i dag mest tilläm-

II Fakta 1

Beräkning av responsivitet

Responsivitet kan beräknas utifrån distributionsparametrar. Det innebär exempelvis att en skillnad i medeltal mellan mätillfällen (t ex mellan baslinje och uppföljning) divideras med ett spridningsmått (t ex standardavvikelsen vid baslinjen): effect size. Det kan också innebära skillnaden mellan två mätillfällen dividerad med standardavvikelsen för denna skillnad: standardized response mean. Tumregler för värdering finns. Till exempel indikerar en kvot under 0,20 trivial effekt och en kvot över 0,80 stor effekt.

pade generella effektmåttet, närmast en »gylle standard«, vilket underlättar tolkning av resultat. Patientgrupp och frågeställningar/hypoteser faller slutligt avgörande vad gäller bästa val av sjukdomsspecifika mätinstrument eller tillägg i kommande studier.

Kvalitetsbedömning av livskvalitetsmått

Medical Outcomes Trust (MOT) är en ideell organisation som garanterar tillgängligheten av goda resultatmått vad gäller hälsostatus och hälsorelaterad livskvalitet samtidigt som standardiseringen av innehåll, poängsättning, skalkonstruktioner och benämning skyddas. Syftet har varit att åstadkomma universell spridning av mätinstrument som uppfyller särskilda krav på psykometrisk kvalitet, definierade av en internationellt sammansatt expertgrupp 1995 [2]. Konstruktörer av dylika resultatmått inviteras att ansöka om MOT-auktorisering. Verksamheten har skärpt uppmärksamheten så att standardiserade mått ökat i användning, även om ytterligare riktlinjer behövs för tolkning av resultat från kliniska utvärderingsförsök/kliniska prövningar. Flera initiativgrupper, t ex European Regulatory Issues Quality of Life Assessment (ERIQA) Group [3] och Cochrane Health-Related Quality of Life Methods Group (HRQL MG) [4], verkar för mer enhetlig dokumentation inom området.

Bedömning av effekttorlekar

Ett nyckelbegrepp bland MOT-kriterierna är responsivitet, dvs ett mätinstrumentets förmåga att reflektera förändringar som kan vara kliniskt viktiga. Olika sätt att beräkna effekttorlekar har därvid använts (Fakta 1).

Metoderna är självklart generella och har redovisats i forskningssammanhang sedan mitten av 1980-talet, bl a inom reumatologisk forskning. Förändringar i formulärdata har därvid också jämförts med förändringar i traditionella kliniska mått. Fördelen är att man kan bli oberoende av mätinstrumenten vid jämförelse av effekter. Nackdelen är att formulärdata strikt sett aldrig uppfyller kraven på intervallskala utan har olika lätt att förändras utefter skalans sträckning. Det är således inte säkert att man kan jämföra en förändring från t ex 50 till 60 poäng på en 0–100-skala med en förändring från 10 till 20. Nedan ges några exempel på kritiska frågor som följer:

- Är det rättvisande att jämföra en förändring i ett livskvalitetsinstrument med en förändring av blodgasmätningar eller spirometriverden?
- Är det rimligt att jämföra storleken på en förändring i ett livskvalitetsformulärs fysiskt orienterade skalor med storleken på en förändring i psykosocialt orienterade skalor?
- Hur ska man jämföra effekttorlek inom skalans »normala nivåer« med effekttorlek inom patologiska nivåer, t ex

sänkt stämningsläge versus depressivitet/klinisk depression?

Beräkning av effekttorlekar bör komplettera, inte ersätta, annan konventionell statistik inklusive hypotesprövning och utnämning av primära och sekundära resultatmått [5].

Exemplet astma/kroniskt obstruktiv lungsjukdom

Figur 1 visar begrepp och instrument längs ett tänkt kontinuum enligt ovan för att klargöra innehåll och position hos vanligen förekommande resultatmått vid astma och kroniskt obstruktiv lungsjukdom [1]. Fler specifika och generella mått från samma rapport finns sammanställda i Fakta 2.

Kliniskt meningsfull förändring. Förutom god effekttorlek enligt ovan är klinisk relevans väsentlig, dvs att det rör sig om s k ankar- eller kriteriebaserad förändring. Exempelvis har på effekttorleksgrund fem poängs skillnad i SF-36, 0–100-skala, av upphovsmännen framförts representera ett tröskelvärde för en kliniskt och socialt relevant förändring [6]. Denna robusta bedömning utifrån ett begränsat antal kliniska studier, där patienter med sjukdom av olika svårighetsgrad kunnat urskiljas väl genom hälsoprofilen, har senare fått stöd från ett stort antal kliniska prövningar. Kön- och åldersperspektivet ska särskilt beaktas. En generell trend att kvinnor rapporterar mer symtom, funktionsinskränkningar och illabefinnande än män är välkänd från epidemiologisk och klinisk forskning. Likaså kan en långsam och kontinuerlig försämring ses i fysiskt orienterade skalor med stigande ålder medan psykosociala aspekter av livskvaliteten blir sämre först efter 70-årsåldern. Att vara fränksild, ha låg utbildningsnivå, vara arbetslös eller bo i storstadsregion är exempel på sociala karakteristika med negativ inverkan på livskvalitetsmått.

Flera specifika mått vid astma och kroniskt obstruktiv lungsjukdom har genomgått särskilt upplagda studier med enda mål att fastställa tröskelvärden för kliniskt meningsfull förändring. Både patienter och vårdgivare har rekryterats och fått värdera det principiellt viktiga: vid vilken minsta poängskillnad i respektive formulärs skala en förändring i hälsotillståndet kan upplevas eller iakttas. Denna forskning har t ex visat att 4 poängs skillnad, 0–100-skala, i SGRQ och 0,5 poängs skillnad, 1–7-skala, i CRQ och AQLQ utgör minsta skillnad av betydelse (minimal important difference, MID) (se Figur 1 och Fakta 2 för tolkning av instrumentförkortningar). Dessa resultat gäller vuxna. Formulär för barn och föräldrar, främst PAQLQ och PCAQLQ, har nyligen börjat användas, och kliniskt meningsfull förändring återstår att fastställa.

Kliniskt underbyggda tröskelvärden kan tillåta meningsfulla beräkningar av antalet patienter som behöver behandlas för att en patient ska nå framgång (number-needed-to-treat, NNT). Sådan tolkning av livskvalitetsdata har nyligen visats ha kompletterande värde vid en randomiserad och kontrollerad korsstudie (cross-over) av salmeterol vs salbutamol vs placebo vid astma. NNT-analysen av tidigare publicerade data [7] visade att fyra patienter behövde behandlas med salmeterol för att en patient skulle få kliniskt meningsfull förbättring över den som åstadkommit med salbutamol. Motsvarande NNT för salmeterol vs placebo var två patienter [8].

Hälsorelaterad livskvalitet var ett av de fyra effektmått som låg till grund för urval av studier och slutsatser av behandling i SBU-rapporten [1]. De studier som refereras till i det följande finns granskade i rapporten. De är randomiserade kontrollerade studier eller kliniska kontrollerade studier och uppfyller höga krav på evidens.

Kortverkande beta-2-stimulerare och antikolinergika. Effekten på livskvalitet är inte studerad. En anledning kan vara att des-

Begrepp Sjukdomsrelaterade/generella	Instrument Respiratoriska/generella																								
Tillståndsspecifika Symtom/besvär/konsekvenser	<table border="0"> <tr> <td>SGRQ</td> <td>CRQ</td> <td>AQLQ</td> <td>LWAQ</td> </tr> <tr> <td>• Symtom</td> <td>• Dyspné</td> <td>• Symtom</td> <td>• Fysisk hälsa</td> </tr> <tr> <td>• Aktivitet</td> <td>• Trötthet</td> <td>• Aktivitet</td> <td>• Psykisk hälsa</td> </tr> <tr> <td>• »Impacts«</td> <td>• Emotionell</td> <td>• Omgivning</td> <td>• Total</td> </tr> <tr> <td>• Total</td> <td>• »Mastery«</td> <td>• Emotionell</td> <td>(11 skalor)</td> </tr> <tr> <td></td> <td>• Total</td> <td>• Total</td> <td></td> </tr> </table>	SGRQ	CRQ	AQLQ	LWAQ	• Symtom	• Dyspné	• Symtom	• Fysisk hälsa	• Aktivitet	• Trötthet	• Aktivitet	• Psykisk hälsa	• »Impacts«	• Emotionell	• Omgivning	• Total	• Total	• »Mastery«	• Emotionell	(11 skalor)		• Total	• Total	
SGRQ	CRQ	AQLQ	LWAQ																						
• Symtom	• Dyspné	• Symtom	• Fysisk hälsa																						
• Aktivitet	• Trötthet	• Aktivitet	• Psykisk hälsa																						
• »Impacts«	• Emotionell	• Omgivning	• Total																						
• Total	• »Mastery«	• Emotionell	(11 skalor)																						
	• Total	• Total																							
Funktion: generella Fysiska/ rörlighetsorienterade konsekvenser	<p>SIP</p> <ul style="list-style-type: none"> • Total (12 kategorier) • Fysisk dimension (3 kategorier): gång, rörlighet, personlig skötsel • Fristående (5 kategorier): hushållsarbete, sömn–vila, fritid–rekreation, arbete, födointag 																								
Fysisk hälsa	<p>SF-36</p> <ul style="list-style-type: none"> • Fysisk hälsokomponent (4 skalor): fysisk funktion, rollfunktion–fysiska begränsningar, smärta, allmän hälsa 																								
Sociala/emotionella/kognitiva konsekvenser	<p>SIP</p> <ul style="list-style-type: none"> • Psykosocial dimension (4 kategorier): social interaktion, psykisk balans, uppmärksamhet–koncentration, kommunikation 																								
Psykisk hälsa/välbefinnande: generella	<p>SF-36</p> <ul style="list-style-type: none"> • Psykisk hälsokomponent (4 skalor): vitalitet, social funktion, rollfunktion–känslomässiga begränsningar, psykiskt välbefinnande <p>HAD</p> <ul style="list-style-type: none"> • Depression • Ångest 																								
Livskvalitet totalt sett	<ul style="list-style-type: none"> • Globalskattningar, globalindex 																								

Figur 1. Hälsorelaterad livskvalitet – kontinuum av begrepp och vanliga instrument vid astma/KOL. (Modifierat från Sullivan och medarbetare [17]).

sa preparat varit väl etablerade långt innan livskvalitetsstudier blev aktuella.

Inhalationssteroider. Flera studier har visat att inhalationssteroider ökar livskvaliteten. Såväl generella som specifika frågeformulär har använts. Patienter med svår, oralsteroidberoende astma som behandlas med inhalationssteroider (flutikason) får förbättrad livskvalitet mätt med SF-36 [9]. Det är i första hand begränsningar i fysisk funktion och begränsningar i vanliga regelbundna aktiviteter som minskar, men uppfattningen om hälsan i allmänhet och det psykiska välbefinnandet ökar också. Vid jämförelse av effekterna av budesonid i dosen 200 µg respektive 400 µg en gång per dag förbättras livskvaliteten mätt med AQLQ i båda budesonidgrupperna jämfört med placebo [10]. Författarna redovisade dock inte

vilken eller vilka domäner som förbättrades mest. I en studie av patienter med medelsvår astma, där man jämförde 100, 200 och 500 µg flutikasonpropionat i pulverform med placebo, användes SF-36 [11]. I alla grupperna visade de två ovan nämnda fysiska SF-36-variablerna förbättring. Vid de två högsta doserna förbättrades psykiskt välbefinnande, vitalitet och summaindex för mental hälsa.

Inhalationssteroid och långverkande beta-2-stimulerare. Endast två studier är publicerade där man jämfört effekten på livskvalitet av beklometason och salmeterol med budesonid enbart respektive med placebo [12, 13]. LWAQ användes. Man fann ingen skillnad mellan grupperna.

Teofyllin. Endast en studie av livskvalitet och teofyllinbe-

II Fakta 2

Hälsorelaterad livskvalitet – vanliga specifika och generella mått vid astma/kroniskt obstruktiv lungsjukdom (KOL)

Instrumentförkortningar för standardformulär i Figur 1

Tillståndsspecifika: astma och KOL

SGRQ: St George's Respiratory Questionnaire [18]
Svensk version [19]

Tillståndsspecifika: KOL

CRQ: Chronic Respiratory Questionnaire [20]
(svensk version under utveckling)

Tillståndsspecifika: astma

AQLQ¹: Asthma Quality of Life Questionnaire [21]
(svensk version föreligger)

AQLQ(S): Standardized version of the AQLQ [22]

LWAQ: Living With Asthma Questionnaire [23]
(svensk version föreligger)

Generella: funktion

SIP¹: Sickness Impact Profile [24]
Svensk version [25]

Generella: funktion & välbefinnande

SF-36¹: Short Form-36 Health Survey [26]
Svensk version [27]

Generella: psykisk hälsa

HAD: Hospital Anxiety and Depression scale [28]
Svensk version [29]

Instrumentförkortningar för standardformulär, ej i Figur 1

Tillståndsspecifika: astma

PAQLQ¹: Paediatric Asthma Quality of Life Questionnaire [30]
Svensk version [31]

PCAQLQ¹: Paediatric Caregiver's Asthma Quality of Life Questionnaire [32]
Svensk version [33]

CAQC: Childhood Asthma Questionnaire Form C [34]

AQLQ: Asthma Quality of Life Questionnaire [35]

Generella: funktion

NHP: Nottingham Health Profile [36]
Svensk version [37]

Generella: funktion & välbefinnande

MHIQ: McMaster Health Index Questionnaire [38]

Generella: psykisk hälsa

GHQ: General Health Questionnaire [39]

POMS: Profile of Mood States [40]

Instrumentförkortningar för globala (hälsoekonomiska) index

QWB: Quality of Well-Being scale [41]

EQ-5D: EuroQol [42]; inkluderar svensk version

HUI: Health Utilities Index [43]

15D: Fifteen-dimensional measure of health-related quality of life [44]

¹Auktoriserat av Medical Outcomes Trust [2]

handling är publicerad [14]. Den avsåg att jämföra effekten av salmeterol med oralt teofyllin. Man fann att båda behandlingarna förbättrade livskvaliteten. Det förelåg däremot inte någon skillnad mellan grupperna. En italiensk version av LWAQ användes.

Antileukotriener. Samtliga fyra domäner i AQLQ visade förbättring när patienterna i en placebokontrollerad multicenter-

studie av stabil astma behandlades med montelukast [15]. AQLQ har också använts i en studie som jämfört montelukast, budesonid och placebo [16]. Alla fyra domänerna visade förbättring för båda behandlingarna jämfört med placebo.

Hyposensibilisering. Förändringar i livskvalitet vid hyposensibilisering på indikationen astma är inte studerade.

Sammanfattning: exemplet astma

Att utvärdera effekten av astmabehandling på individens livskvalitet, dvs en effekt ur patientens synvinkel, är väsentligt. Det finns i dag mätinstrument tillgängliga som uppfyller höga vetenskapliga krav. Dessa instrument har utvecklats de senaste årtiondena. Därför är det först de senaste åren som livskvalitetsmätningar vid farmakologiska studier har blivit vanliga. Frånvaron av bevis för att ett visst astmaläkemedel påverkar livskvaliteten behöver således inte betyda att effekt saknas – det kan helt enkelt vara så att det inte är studerat. Detta gäller exempelvis etablerade läkemedel såsom beta-2-stimulerare och antikolinergika.

Livskvalitet ett bra effektmått i kliniska studier

Under de senaste decennierna har tillförlitlig metodik för mätning av livskvalitet i sjukvården utvecklats och prövats i internationellt samarbete mellan läkare och metodexperter. Etablerade livskvalitetsmått i sjukvården är alltid multidimensionella, kvantitativa och uppbyggda enligt psykometrisk teori till sammansatta skalor, profiler och index med känd reliabilitet, validitet och känslighet för kliniskt relevanta förändringar. Dessa formulär, som patienten själv kan hantera med kortfattad vägledning, omfattar väldefinierade nyckelområden som symtombörda, funktionsinskränkningar och välbefinnande fysiskt, psykiskt och socialt. Metodiken har värderats så att riktlinjer för användning och tolkning kunnat anges. De vanligast förekommande resultatmått har anpassats för och prövats under svenska förhållanden. Viktiga fördelar med livskvalitet som effektmått är:

- patientbaserad informationskälla, självskattning
- mångdimensionellt begrepp, hälsorelaterat
- god förmåga att styrka förändringar i hur patienten mår och fungerar i vardagen
- kan ge vägledning i dimensioneringen av behandlingsstudier.

Livskvalitetsmått har en framtid i vår sjukvård

Livskvalitetsdata, som belyser behandlingsresultat för grupper av patienter, har redan bidragit till förbättringar i vården av den enskilde åtminstone på ett generellt plan. Dokumentation av forskning på individnivå har just startat, men klart är att behandlingsregimer kan korrigeras på grundval av livskvalitetsdata så att patienternas motivation att påbörja och fullfölja behandling ökar och så att effekterna optimeras. Vårdtagaren själv kan få hjälp att bättre lägga märke till viktiga symtom och ges möjlighet att informera vårdgivaren bättre om sjukdomsbörda och bieffekter av behandling, vilket underlättar att en skräddarsydd behandling kommer till stånd. För vårdgivaren ger mätningarna ny information om patienten. Ansvariga för policybeslut i vården kan få vägledning av patientbaserad information vid t ex uppläggning av nya vårdprogram.

Referenser

1. Behandling av astma och KOL. En systematisk kunskapsställning. Stockholm: SBU, Statens beredning för medicinsk utvärdering; 2000. Rapport nr 151.
2. The Medical Outcomes Trust: Scientific Advisory Committee. In-

- strument review criteria. *MOT Bulletin* 1995;3(4):1-4.
3. Acquadro C for the ERIQA Group. The European Regulatory Issues on Health-Related Quality of Life Assessment (ERIQA) Project: Preliminary results and plans for the future. *Quality of Life Newsletter* 2000;24:1-3.
 4. Staniek V. Cochrane Health-Related Quality of Life Methods Group (HRQoL MG): toward registration. *Quality of Life Newsletter* 2001;26:1-2.
 5. Guyatt GH, King DR, Feeny DH, Stubbing D, Goldstein RS. Generic and specific measurement of health-related quality of life in a clinical trial of respiratory rehabilitation. *J Clin Epidemiol* 1999;52:187-92.
 6. Ware JE Jr, Snow KK, Kosinski M, Gandek B. SF-36 Health survey manual and interpretation guide. Boston: New England Medical Center: The Health Institute; 1993.
 7. Juniper EF, Johnston PR, Borkhoff CM, Guyatt GH, Boulet LP, Haukioja A. Quality of life in asthma clinical trials: comparison of salmeterol and salbutamol. *Am J Respir Crit Care Med* 1995;151:66-70.
 8. Juniper EF. Interpreting of quality of life data. *Quality of Life Newsletter* 1999;23:3.
 9. Noonan M, Chervinsky P, Busse WW, Weisberg SC, Pinnas J, de Boisblanc BP, et al. Fluticasone propionate reduces oral prednisone use while it improves asthma control and quality of life. *Am J Respir Crit Care Med* 1995;152:1467-73.
 10. McFadden ER, Casale TB, Kemp JP, Metzger WJ, Nelson HS, et al. Administration of budesonide once daily by means of turbuhaler to subjects with stable asthma. *J Allergy Clin Immunol* 1999;104:46-52.
 11. Mahajan P, Okamoto LJ, Schaberg A, Kellerman D, Schoenwetter WF. Impact of fluticasone propionate powder on health-related quality of life in patients with moderate asthma. *J Asthma* 1997;34:227-34.
 12. Hyland ME, Crocker GR. Validation of an asthma quality of life diary in a clinical trial. *Thorax* 1995;50:724-30.
 13. Vermetten FA, Boermans AJ, Luiten WD, Mulder PG, Vermue NA. Comparison of salmeterol with beclomethasone in adult patients with mild persistent asthma who are already on low-dose inhaled steroids. *J Asthma* 1999;36:97-106.
 14. Nutini S, Martini T, Righi R. Long-term treatment of asthmatic patients with salmeterol vs slow-release theophylline. *Respir Med* 1998;92:683-90.
 15. Reiss TF, Chervinsky P, Dockhorn RJ, Shingo S, Seidenberg B, Edwards TB. Montelukast, a once-daily leukotriene receptor antagonist, in the treatment of chronic asthma: a multicenter, randomized, double-blind trial. Montelukast Clinical Research Study Group. *Arch Intern Med* 1998;158:1213-20.
 16. Malmström K, Rodriguez-Gomez G, Guerra J, Villaran C, Pineiro A, Wei LX, et al. Oral montelukast, inhaled beclomethasone, and placebo for chronic asthma. A randomized, controlled trial. Montelukast/Beclomethasone Study Group. *Ann Intern Med* 1999;130:487-95.
 17. Sullivan M, Karlsson J, Taft C. How to assess quality of life in medicine: Rationale and methods. In: Guy-Grand B, Ailhaud A, editors. *Progress in obesity research*: 8. London: John Libbey; 1999. p. 749-55.

I Läkartidningens elektroniska arkiv
<http://tarkiv.lakartidningen.se>
 är artikeln kompletterad med fullständig referenslista.

SUMMARY

Health-related quality of life informative as outcome measure in clinical studies
 Practical example: asthma/chronic obstructive pulmonary disease

Marianne Sullivan, Alf Tunsäter

Läkartidningen 2001;98:4428-33

Methods for measuring health-related quality of life have been developed and proven valid during the past two decades. It was accomplished through international collaborations between clinicians and method experts. Standardized questionnaires, self-administered by patients, include key domains such as symptoms, functional limitations and well-being from a physical, mental and social perspective. The commonly used instruments are adapted and validated for Swedish conditions. Clinical trials can be powered also according to these outcome measures, as they show documented responsiveness to important changes in patients' everyday functioning and well-being. Patient-reported outcome of treatment efficacy is nowadays adequately standardized for scientific reviews according to the conventional criteria of evidence based medicine. Most importantly, health-related quality of life measures will provide new information of clinical value as demonstrated here in the example of asthma/chronic obstructive pulmonary disease. Thus, outcome can be better assessed by changes in the total burden of disease and its consequences than in terms of pulmonary function tests or biochemical markers.

Correspondence: Marianne Sullivan, Health Care Research Unit, Dept of Internal Medicine, Sahlgrenska universitetssjukhuset, SE-413 45 Göteborg, Sweden
 (marianne.sullivan@medicine.gu.se)

Bengt Järholm, professor vid institutionen för folkhälsa och klinisk medicin, Umeå universitet, överläkare vid yrkes- och miljömedicinska kliniken, Norrlands Universitetssjukhus, Umeå
(bengt.jarvholm@envmed.umu.se)

En observationsstudie kräver grundlig eftertanke av läsaren

■ Mycket kunskap om orsaker till sjukdomar grundas på observationsstudier. Sådana brukar ibland ställas i motsats till experimentella studier. I de senare bestämmer försöksledaren villkoren, vanligen genom en randomisering till den eller de faktorer som hon avser att studera. En annan uppdelning som är vanlig är den mellan kvantitativa och kvalitativa studier. I denna artikel behandlas kvantitativa observationsstudier.

Det typiska för observationsstudien är att forskaren studerar fenomen under betingelser som han/hon inte fullständigt styr över. Genom mer eller mindre skickliga metoder kan dock de problem som detta medför bemästras. Metodiken vid observationsstudier beskrivs i läroböcker i epidemiologi [1]. Syftet med denna artikel är dock inte att lära läsaren att göra en observationsstudie utan att som observant läsare förstå dess svagheter och förtjänster. Det finns flera böcker som beskriver hur läsaren kritiskt kan analysera en sådan studie [2-4].

När man ska läsa en observationsstudie kan det ibland hjälpa att jämföra den med en experimentell studie. Vissa teoretiker inom området har tagit den experimentella studien som modell, och menar att när man utformar en observationsstudie ska man försöka få den att så långt som möjligt efterlikna en experimentell studie [5]. Det centrala i en experimentell studie är att man har en studiepopulation, vilken man randomiserar i en grupp som utsätts för en faktor och en grupp som får fungera som kontroll.

Observationsstudier har sin styrka i sådana situationer där en experimentell studie inte låter sig genomföras. Det går till exempel aldrig att göra en studie av rökningens skadliga effekter hos människa genom att låta en slumpmässigt utvald grupp röka från 15 till 55 års ålder och en annan grupp bara röka från 35 års ålder. Det finns till exempel ingen kunskap om riskerna med att börja röka vid 50 års ålder eftersom ingen gör det. Det skulle kunna vara farligare eller mindre farligt än att börja röka som ung. Uppfattningen att det är farligare att börja röka som mycket ung bygger på studier av personer som börjat röka i åldrarna 10–25 år.

Det bör påpekas att i situationer där det går att göra randomiserade kontrollerade studier, till exempel prövning av läkemedel eller operationsmetoder, är sådana att föredra. I en observationsstudie försöker man justera för kända störfaktorer och minimera andra typer av bias såsom selektionsfel (se

SAMMANFATTAT

En observationsstudie kräver många överväganden av sin läsare.

Observationsstudien fyller sin funktion i situationer när det inte är möjligt att göra randomiserade studier.

Observationsstudien ger en bra möjlighet att uppskatta risken i verkliga situationer.

Bias kan göra att risker både över- och undervärderas. Observationsstudier är ofta en viktig länk när det gäller att bedöma kausalitet.

Evidensbaserad medicin

Fakta 1 för innebörd av vissa termer som används i artikeln). I en tillräckligt stor randomiserad studie gör slumpen att även okända störfaktorer fördelas jämnt mellan grupperna och att det är osannolikt att selektionsfel uppträder.

Som läsare av en studie är det bra att göra klart för sig studiens uppläggning och vilka jämförelser som görs. Många är imponerade av stora studier. Jag ska försöka övertyga läsarna av denna artikel om att en stor observationsstudie inte behöver innebära en studie av hög kvalitet. Eftersom observationsstudier ofta används i diskussioner kring kausalitet går det knappast att undvika att också diskutera deras roll när man ska bedöma kausalitet. Det går också att ställa upp några punkter att särskilt tänka på när man läser en positiv respektive negativ observationsstudie, vilket diskuteras till sist.

Uppläggning

Principiellt finns det två sätt att göra observationsstudier. Antingen utgår undersökaren från två grupper där den ena gruppen, av omständigheter som undersökaren inte kontrollerar, utsätts för den faktor som ska studeras (till exempel tobaksrök eller högt kolesterol) och den andra gruppen (referenter-

II Fakta 1

Innebörd av termer som förekommer i artikeln

Bias: Samlingsterm för alla typer av icke slumpmässiga fel i en studie, till exempel störfaktorer, fel i mätning av exponering och selektionsfel.

Störfaktor: Faktor som är vanligare i en av de undersökta grupperna och som också kan orsaka den sjukdom som studeras.

Negativ studie: Undersökning som visar att det inte föreligger ett samband mellan en faktor och en sjukdom.

Positiv studie: Undersökning som påvisar samband mellan en faktor och en sjukdom.

Extremvärde (outlier): Kan vara en kombination av flera mätta variabler som är ovanlig i gruppen. Det är särskilt viktigt att vara observant på extremvärden i studier som använder till exempel multivariata statistiska tekniker, där hela »sambandet« kan bygga på ett eller ett fåtal extremvärden.

Reststörfaktor: Undersökningen har försökt ta hänsyn till en störfaktor, till exempel genom regressionsanalys, men metoden ger endast delvis justering för störfaktorn.

na) inte utsätts för faktorn. Sedan undersöks vilka som utvecklar sjukdom under en viss tidsperiod. Så genomfördes till exempel Dolls undersökning om sambandet mellan lungcancer och rökning. Där tillfrågades vid ett tillfälle läkare om de rökte eller inte, och sedan studerades vilka som därefter drabbades av lungcancer. Denna uppläggningsmetod är lätt att förstå och kallas vanligen för kohortstudie. Den liknar också den experimentella studien, dock med skillnaden att försökspersonerna inte randomiseras till de två grupperna.

Det andra sättet är att bland sjuka undersöka hur vanlig en viss faktor är, och sedan hos en kontrollgrupp studera hur vanlig faktorn är. Detta brukar kallas fall-kontrollstudie. Ett exempel på en sådan är studierna av sambandet mellan bekämpningsmedel och mjukdelstumörer, som bland annat genomförts i Sverige [6]. Undersökarna frågade personer som drabbats av en sådan tumör om de varit utsatta för bekämpningsmedel. Samtidigt frågade man personer slumpmässigt utvalda ur befolkningen om de varit utsatta för bekämpningsmedel. De som drabbades av tumörer var oftare utsatta för bekämpningsmedel, vilket tyder på att det finns ett samband mellan sådana tumörer och bekämpningsmedel.

Det finns a priori ingen kvalitetsskillnad mellan en kohortstudie och en fall-kontrollstudie från teoretisk synpunkt. Båda har sina styrkor och svagheter. Allt beror på hur väl upplagda och genomförda studierna är. En svårighet med fall-kontrollstudien är att den ibland är svår att begripa för läsaren. Detta gäller särskilt matchade fall-kontrollstudier, som innebär en stor utmaning både för läsaren och för den som gör en sådan. För en utförligare diskussion av detta hänvisas till läroböcker i epidemiologi.

En svaghet som ofta framhålls hos fall-kontrollstudien är risken att den som är sjuk försöker komma ihåg vad som hänt, medan den som är kontrollperson och kanske frisk inte tänker efter så noga, så kallad »recall bias«. Det har till exempel i ovan refererade studier av bekämpningsmedel och cancer diskuterats om patienter som drabbats av tumörer ansträngde sig mycket mer att komma ihåg eventuell exponering för be-

kämpningsmedel än personer som slumpmässigt valdes ur befolkningen. Detta problem kan begränsas genom att använda andra metoder för att avgöra exponeringen hos fall och kontroller. I en undersökning av leukemi hos sjömän byggde exponeringsuppgifterna på data hämtade ur ett register över sjömän och fartyg, vilket naturligtvis inte är känsligt för »recall bias« [7].

En speciell variant av dessa studier är tvärsnittsstudien, där både de faktorer som tros orsaka en viss sjukdom och sjukdomen studeras samtidigt. Det kan då vara svårt att reda ut vad som är orsak och verkan. Om sjuklighet är högre hos arbetslösa än hos dem som arbetar kan det vara svårt att avgöra om detta beror på att man blir sjuk av att vara arbetslös eller om de som är sjuka oftare blir arbetslösa. Ibland är det dock mindre problematiskt. Många genetiska studier undersöker till exempel hur vanlig en viss gen är hos sjuka respektive i normalbefolkningen.

Studiens interna validitet

Eftersom det inte sker en randomisering av personer till den exponerade gruppen respektive kontrollgruppen kan dessa grupper skilja sig på andra sätt än genom den faktor som ska studeras. Sådana systematiska fel, bias, finns i de flesta studier. För läsaren gäller det framför allt att försöka bedöma hur betydelsefulla de är och deras riktning; det vill säga underskattas risken eller överskattas risken på grund av bias? Vanligen brukar man skilja mellan tre typer av bias som beskrivs nedan.

Störfaktorer: En av de luftföroreningar som diskuteras mycket idag är dieselavgaser. Det finns stort intresse kring frågan om huruvida exponering för dieselavgaser innebär en ökad lungcancerriks, och i så fall hur stor den är. Ett flertal observationsstudier har noterat att lastbilschaufförer har en ökad risk för lungcancer, vilket tolkats bero på deras exponering för dieselavgaser [8, 9]. Lastbilschaufförernas incidens av lungcancer jämförs med allmänpopulationens risk eller risken i någon annan yrkesgrupp. I sådana jämförelser krävs naturligtvis att man först tar hänsyn till ålder, eftersom ålder har ett starkt samband med risken för lungcancer. Det finns idag olika tekniker att göra detta, och i allmänhet kan det göras med hög precision. Det är sedan tidigare känt att tobaksrökning, men också vissa kostfaktorer liksom om personen bor i radonhus eller inte, kan ha betydelse för uppkomst av lungcancer. Måttligt stora skillnader i rökvanor mellan två grupper kan resultera i måttligt stora skillnader i lungcancerriks mellan grupperna. Det skulle också kunna finnas kostfaktorer som skiljer mellan grupperna. Lastbilschaufförer skulle till exempel kunna äta mycket mindre grönsaker än kontrollgruppen, och det skulle kunna förklara en viss överrisk.

Rökning utgör i exemplet ovan en faktor som kan »störa« analysen av sambandet mellan dieselavgaser och lungcancer. En sådan faktor brukar kallas för en störfaktor (confounder). För att rökning ska vara en störfaktor måste två villkor vara uppfyllda. Dels måste rökning kunna orsaka den sjukdom som ska studeras, i detta fall lungcancer, dels måste rökvanorna vara olika i de grupper som jämförs. Om till exempel studien avsett malignt melanom hade inte tobaksrökning varit en störfaktor eftersom inte tobaksrökning orsakar malignt melanom. Däremot kunde då exponering för UV-strålning ha varit en störfaktor. Om rökvanorna varit identiska bland lastbilschaufförer och kontroller hade rökning inte varit en störfaktor. En störfaktor kan leda till att samband över- eller underskattas.

Ett sätt att ta hänsyn till potentiella störfaktorer är att man försöker mäta ålder, rökvanor, kostfaktorer etc så långt det går, och sedan justera för skillnader i den epidemiologiska

och statistiska analysen. En annan teknik är att begränsa studien till en viss grupp, till exempel icke rökare eller rökare. För att underlätta jämförelser görs till exempel studier på läkare, sjuksköterskor och studenter då betydelsen av till exempel kroppsvikt, blodtryck och rökvanor för uppkomst av sjukdomar studeras. Genom att begränsa studien till homogena grupper med avseende på potentiella störfaktorer minskas risken för bias.

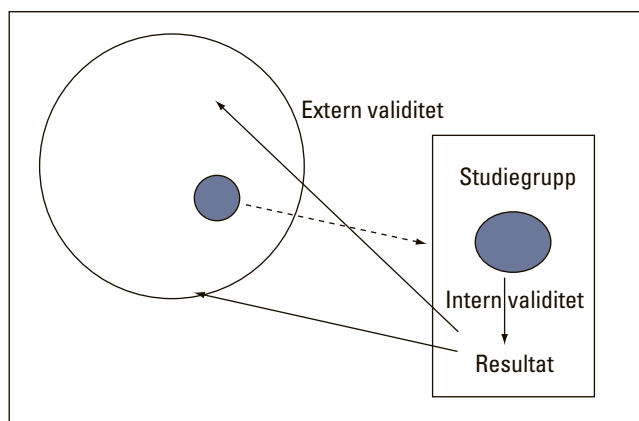
För att man samtidigt ska kunna ta hänsyn till flera störfaktorer brukar idag ofta multivariata statistiska tekniker tillämpas. Man ska dock inte låta sig imponeras alltför mycket som läsare av avancerad statistik. Även om man använder så kallad logistisk regression, poissonregression eller liknande metoder kan det mycket väl finnas kvar olikheter i jämförelsen. Det är inte sällan författaren säger sig ha tagit hänsyn till rökvanor i analysen, men om till exempel rökvanorna klassificeras lite fel eller de statistiska modeller som används (exempelvis poisson- eller logistisk regression) inte helt korrekt beskriver sambanden, kan skillnader i risk fortfarande bero på skillnader i rökvanor. Det brukar på fackspråk kallas »reststörfaktorer«. Om lastbilschaufförerna och kontrollerna i exemplet ovan klassificerades i icke-rökare, ex-rökare och rökare, och hänsyn togs till eventuella skillnader i en poissonanalys, kan det fortfarande finnas kvar störfaktorer av rökning. Lastbilschaufförer kan till exempel i genomsnitt röka 20 cigaretter per dag, medan rökande kontroller bara röker i genomsnitt 10 cigaretter per dag.

Mätfel: Det finns nästan alltid mätfel vid till exempel uppskattning av rökvanorna ovan. Om felet är olika mellan grupperna kan bias uppstå. Flera studier av passiv rökning har gått till så att risken för lungcancer hos icke-rökare har jämförts mellan dem vars make/maka rökt respektive inte rökt. Det är lätt att inse att det kanske är något större risk att en rökare gifter sig med en rökare och en icke-rökare med en icke-rökare. En viss felklassificering av icke-rökare är inte helt osannolik, det vill säga enstaka personer som uppger sig vara icke-rökare kan till exempel ha rökt tidigare. Då det är troligare att rökare lever ihop med rökare än med icke-rökare blir resultatet förhöjda risker för lungcancer hos »icke-rökare« som lever tillsammans med rökare, vilket dock egentligen beror på en felklassificering av icke-rökarna [10]. Det är dock knappast troligt att det förklarar hela överrisken i lungcancer bland personer utsatta för passiv tobaksrökning [11].

I en aktuell studie av lungcancerriksen hos icke rökande svenska byggnadsarbetare fann vi att om analysen genomfördes på personer som vid flera tillfällen uppgett att de var icke-rökare och jämförde den med incidensen bland dem som åtminstone vid ett tillfälle uppgett sig vara icke-rökare hade de som vid flera tillfällen uppgett att de var icke-rökare en lägre risk för lungcancer [12]. Denna skillnad tolkar vi bero på att frågar man flera gånger minskar risken för att en person som tidigare rökt säger att han aldrig rökt.

Om mätfelen är slumpmässiga och lika mellan grupperna blir resultatet i allmänhet att det är svårt att påvisa skillnader mellan grupperna eller att skillnaderna underskattas. I det avseendet skiljer sig inte observationsstudien från den randomiserade kontrollerade studien.

Skillnader i urval – selektionsfel: Ett annat fel i jämförelsen kan vara skillnader i urval av populationer, och sådana fel kan vara mycket svåra att justera för i analysen. Om dödligheten i en yrkesverksam grupp jämförs med dödligheten i allmänbefolkningen har i de flesta fall den yrkesverksamma gruppen en något lägre dödlighet. Detta gäller även om hänsyn tas till tänkbara störfaktorer som ålder, kön och rökvanor. Detta tolkas bero på att personer som redan lider av kroniska sjukdo-



Figur 1. Skillnad mellan intern och extern validitet. Bland alla personer som är eller kan tänkas bli utsatta för en faktor studeras bara en begränsad grupp, här skuggad. Om resultaten är giltiga för den studerade gruppen har studien hög intern validitet. Om resultaten dessutom kan generaliseras till många andra grupper (hela den stora cirkeln) har den hög extern validitet. Den lilla cirkeln kan vara män i åldern 40–60 år, medan den stora cirkeln kan bestå av män och kvinnor i alla åldrar.

mar, har missbildningar eller är svåra missbrukare har svårt att få anställning. Om allmänbefolkningen, som ju också består av bland annat kroniskt sjuka och missbrukare, jämförs med personer som har fått anställning är det uppenbart att de som fått anställning vid en viss tidpunkt sannolikt också är något friskare om de undersöks vid en senare tidpunkt. Ibland har dessutom hälsokontroller krävts före anställning, vilket ytterligare bidrar till sådana skillnader. Detta brukar kallas »healthy worker bias«. En sådan bias går inte att korrigera med en statistisk analys. Det blir i stället upp till läsaren av studien att försöka bedöma storleken av ett sådant selektionsfel. Skulle en yrkesverksam grupp ha högre dödlighet än allmänbefolkningen talar det snarast för att risken är ännu större om man använt en helt adekvat kontrollgrupp.

Studiens externa validitet

Om observationsstudien interna validitet är sådan att man vågar lita på resultaten uppstår frågan om man kan generalisera fynden. Omfattar studien endast män är det inte självklart att resultaten också är giltiga för kvinnor. Överväganden som har med generaliserbarhet att göra brukar kallas för extern validitet (Figur 1). Vid bedömningen av den externa validiteten har ofta val av studiepopulation stor betydelse. I detta avseende skiljer sig inte observationsstudien från den randomiserade kontrollerade studien.

Många välkända observationsstudier bygger på hälsokontroller av personer vid ett visst tillfälle, som Framinghamstudien och 1913 års män. Det görs en hälsokontroll, och sedan följs populationen framåt i tiden. Det som sällan diskuteras är att urvalet bara består av 50–80 procent av dem man erbjöd hälsokontrollen. Det finns ofta anledning att tro att de som inte kom var personer med andra sjukdomar, personer med missbruk eller en svår social situation etc. Resultaten från studien kan då inte självklart överföras till att gälla alla människor, kanske särskilt inte sådana som redan är sjuka eller missbrukare.

Även när det gäller studier av genetiska markörer, som oftast är observationsstudier, är det viktigt att noga vara medveten om vilken grupp studien görs. För några år sedan kom flera rapporter om att vissa gener (BRCA1 och BRCA2) förutsäddes en mycket hög risk för bröstcancer. Det som ofta inte framgick, till exempel i medier, var att de första studierna gjordes i vissa familjer där bröstcancer var mycket vanligt.

Möjligheten att förutsäga risk visade sig sedan inte vara lika stor i andra familjära anhopningar av bröstcancer, och generna tycks ha liten betydelse bland bröstcancerfall utan familjär anhopning [13-15].

Det kan också vara så att urvalet i hög grad begränsar möjligheten till slutsatser. I olika länder byggs det upp register dit personer med vissa sjukdomar anmäls, till exempel sjukdomar som har inträffat i arbetet. Studier av samband mellan olika faktorer och sjukdomar kan då innebära särskilda problem. Resultaten från sådana register används ofta för att förklara hur vanliga olika orsaker är, till exempel astma [16]. Sådana jämförelser innebär naturligtvis tolkningssvårigheter. Eftersom fallen i registren baseras på anmälningar blir risken inte bara beroende på sambandet mellan exponeringen och sjukdomen utan också på rapporteringsbenägenhet. Den är utomordentligt svår att värdera och kan variera kraftigt.

Storlek på studie

I en randomiserad studie gäller att om studien är tillräckligt stor är chansen att dra fel slutsats liten. Motsvarande gäller inte alls för en observationsstudie. Ett selektionsfel har lika stor betydelse i en observationsstudie vare sig studien är stor eller liten. Samma sak gäller betydelsen av störfaktorer och andra typer av bias. Ska man studera lungcancerrisken av dieslrök gäller det att de grupper man jämför inte har skillnader i rökvanor alternativt att man kan ta hänsyn till dessa skillnader i sin analys. Finns det skillnader eller är de metoder man använder för att justera för skillnader i rökvanor ofullständiga kvarstår skillnaderna oavsett studiens storlek. Om till exempel den relativa risken egentligen är 1,0, men på grund av skillnader i rökvanor blir 1,2, blir den det i en stor såväl som en liten studie. Skillnaden är att i den stora studien blir den relativa risken 1,2 statistiskt signifikant skild från 1,0 men inte i den lilla. Eftersom det närmast tycks vara en ryggmärgsreflex att klassificera en stor studie som kvalitativt högstående kan man lätt imponeras av den statistiska signifikansen. Det kanske kan vara på sin plats att påpeka att det ofta är svårt i mycket stora studier att samla in uppgifter med samma höga noggrannhet som i små studier. Är det så att man mäter störfaktorfel innebär det att det kan föreligga en så kallad reststörfaktor, trots att man i analysen har använt de mest avancerade statistiska metoder.

Kan man »fuska« i en observationsstudie?

Det är naturligtvis så att direkt falsifikation av data kan förekomma också i observationsstudier, och det kan sällan upptäckas av en läsare utan tillgång till originaldata, och i många fall inte ens då. Något som får mig tveksam är extremt hög deltagarfrekvens (över 90–95 procent) i studier som bygger på till exempel enkäter insamlade i befolkningen. Alla som har genomfört sådana studier vet att det är nästan omöjligt att uppnå så höga frekvenser. I en seriös undersökning brukar en så hög deltagarfrekvens också kommenteras, och undersökaren försöker förklara varför den är så hög.

Det finns också sätt att analysera sina data som kanske inte direkt ohederliga, men visste läsaren om hur förfarandet gått till skulle kanske bedömningen av fynden bli annorlunda. De flesta är medvetna om det som brukar kallas masssignifikans. Presenteras många jämförelser i en artikel och någon är statistiskt signifikant kan det bero på slumpen ändå. Det som ofta inte framgår i artikeln är att materialet ofta analyserats på en rad olika sätt, och man väljer att presentera ett av dem. Vid multivariata metoder, där man samtidigt tar hänsyn till en rad störfaktorer, kan indelning av grupper, val av modell, val av antal faktorer som man tar hänsyn till etc ha betydelse, särskilt om överrisker eller signifikanser är måttligt stora. Även hur extremvärden (»outliers«) hanteras kan ha stor betydelse för tolkning-

en. Genom att välja kan således författaren presentera den modell som passar hans/hennes uppfattning om hur sambanden borde vara. Detta kan delvis genomskådas om data också presenteras i stratifierad form. Det är därför jag som läsare också gärna vill ta del av basala analyser; dessa går inte att påverka på samma sätt som de multivariata analyserna.

En studie kan innehålla öppna frågor som sedan ska kategoriseras innan studien bearbetas. Som läsare får man sällan se frågeformuläret. Överföringen från en öppen fråga till data som ska analyseras statistiskt är alltid ett kritiskt moment. Ofta finns tolkningssvårigheter av svaren, vilket lämnar utrymme för undersökarens tolkningar. Om den som klassificerar sådana öppna frågor också är medveten om studiens syfte och vilken grupp patienten tillhör kan det finnas risk för bias.

Kausalitet

Ibland framförs uppfattningen att en observationsstudie inte medger några definitiva slutsatser om samband därför att den inte är gjord under experimentella betingelser. Den kritiske menar att det är först när vi gör experimentet som vi kan vara riktigt säkra. Det är dock ganska enkelt att från sin egen erfarenhet konstatera att det inte riktigt är så enkelt. Det torde vara ganska få som läser denna artikel som inte tror på att tobaksrökning kan orsaka lungcancer. Den kunskapen accepterades och etablerades väsentligen på resultat från observationsstudier. Det går ju inte att göra en randomiserad human studie av risken för lungcancer på grund av tobaksrökning.

Det finns ett stort antal samband som accepterats i huvudsak baserat på resultat från observationsstudier, till exempel betydelsen av flera olika faktorer för uppkomst av cancer, som tobaksrökning, asbest, vissa kromater, mineralolja, bensen, arsenik och flera kostfaktorer. Den inbördes betydelsen av högt blodtryck, kolesterol och diabetes för uppkomst av hjärtinfarkt baseras likaledes i hög grad på resultat från observationsstudier.

Vissa försöker dela in riskfaktorer kvalitativt, så att det antingen finns en risk eller inte finns en risk. Dock innebär de allra flesta faktorer påtagliga risker vid tillräckligt hög dos, till exempel kemikalier, födoämnen, mediciner och arbetsmiljöfaktorer. Det vi vill veta är något om riskens storlek, och eftersom risken ofta beror inte bara på den enskilda faktorn utan på i vilken omgivning den finns krävs studier under sådana förhållanden. Till observationsstudiens stora fördelar hör att den uppskattar storleken på riskerna i det verkliga livet; det vill säga de resultat som kommer från en välgjord observationsstudie ger en bra bild av hur stor risk en faktor innebär. Tobaksrökning tycks till exempel innebära betydligt mindre risk för hjärtinfarkt i sydeuropeiska länder än i nordeuropeiska, den absoluta risken att få lungcancer i samband med asbest är mycket större för rökaren än för icke-rökaren och så vidare. För att travestera ett populärt ord inom beteendevetenskaper kan man säga att betydelsen av en faktor är kontextberoende även inom medicinen.

Det finns olika kriterier som kan användas för att bedöma om en faktor är kausal eller inte, den kanske mest citerade är Hills kriterier, Fakta 2 [1]. Det jag dock vill betona är att när man diskuterar kausalitet bör man försöka väga in alla typer av kunskap man har, oavsett om det är från experimentella studier eller observationsstudier.

Några tumregler

En observationsstudie kräver betydligt mer av läsaren än en randomiserad studie, eftersom man måste bedöma om de jämförelser som görs verkligen är adekvata, om man har tagit hänsyn till alla tänkbara störfaktorer etc. Som läsare av en observationsstudie brukar jag först fundera på om jag begriper hur studien har gjorts, efter att snabbt ha läst igenom den. För-

II Fakta 2

Hills kriterier som kan användas för att bedöma om ett samband är kausalt eller inte

Styrkan: Mycket högre risk bland dem som utsätts för faktorn talar för samband; det vill säga är den relativa risken hög talar det för ett samband.

Konsistensen: Har man funnit samma eller liknande fynd i många undersökningar gjorda av många undersökare talar det för ett kausalsamband.

Specificiteten: Om sjukdomen förekommer i visst organ, har ett visst patohistologiskt utseende och dessutom förekommer vid en speciell exponering talar det för samband.

Tidsrelation: Exponeringen ska föregå sjukdomen. Hill anger som exempel förekomst av tbc hos arbetare i vissa miljöer; detta skulle kunna bero på att arbetare med tbc söker sig till vissa arbetsmiljöer.

Dos-respons: Har de med hög exponering högre risk än de med låg exponering talar det för orsakssamband.

Trovärdighet: Kan sambandet förstås utifrån existerande biologiska teorier och kända mekanismer ger det stöd för att det råder orsakssamband. Hill varnar dock för kravet att ett observerat samband ska ha biologisk trovärdighet. Trovärdigheten beror ju på den biologiska kunskapen vid tillfället.

Koherens: Stämmer fynden med sjukdomens naturalhistoria? Det förhållandet att man kan finna histologiska förändringar i bronkepitel hos rökare och att man med tjära från tobaksrök kan orsaka hudcancer hos djur ger stöd för att inhalation av tobaksrök kan leda till lungcancer.

Experiment: Leder eliminering av faktorn till att förekomsten av sjukdom minskar talar det för ett orsakssamband.

Analogi: Eftersom vi vet att röda hund respektive Neurosedyn (talidomid) kan orsaka fosterskador har vi lättare att acceptera att även andra virus eller läkemedel kan orsaka fosterskador.

står jag inte det blir jag alltid tveksam. Det går också att fundera över problem i en kohort- respektive fall-kontrollstudie. Betydelsefulla bias är ibland något olika i dessa studier. De viktigaste/vanligaste problemen för kohortstudier är störfaktorer, selektionsfel samt begränsad generaliserbarhet därför att studien bara omfattar en viss grupp. De viktigaste/vanligaste problemen för fall-kontrollstudier är »recall bias« och störfaktorer. Ett ytterligare sätt att som läsare angripa problemet är att fundera över om studien är negativ eller positiv.

Är studien negativ, det vill säga inte visar något samband, bör läsaren noga studera hur den faktor som ska studeras har mätts. I studier av sambandet mellan fettintag i kosten och cancer är det inte ovanligt att individens intag av fett uppskattas i enkäter. Har mätmetoden dålig precision är det lätt att förstå att studien kan bli negativ. Det blir ju närmast slumpmässigt om en individ kommer att hamna i den grupp som har högt respektive lågt fettintag. Är studien liten är det också svårt att påvisa en statistiskt säkerställd risk. Men det gäller alla typer av studier.

Läser man en studie som visar ett positivt samband mel-

lan en faktor och sjukdom finns det anledning att överväga om det finns andra faktorer som skiljer mellan den exponerade gruppen och kontrollgruppen. Finns till exempel mer cancer bland dem som intar mycket fett i kosten kan det finnas anledning att fundera över om det finns ett positivt samband mellan fettintag och tobaksrökning. Om tobaksrökning tycks förklara en ökad förekomst av ont i ryggen kan det bero på att personer som har tunga eller påfrestande arbeten röker i högre grad än personer som har arbeten med en bättre miljö.

Jag tycker också att man bör fundera över om det kan vara så att undersökaren redan i förväg visste om sambandet. Detta gäller särskilt mindre studier av oväntade samband. Anta att en läkare på sin mottagning noterar att ovanligt många patienter med hjärtinfarkt dricker en viss sorts kaffe. Görs sedan en fall-kontrollstudie där man studerar hur stor andel av patienterna som dricker detta kaffe finns troligen ett positivt samband. Den korrekta undersökningstekniken i ett sådant fall är naturligtvis att rapportera den första iakttagelsen som en fallstudie. Sedan görs en ny studie där andra patienter ingår.

Som läsare av en observationsstudie menar jag att det oftast är mer intressant att fundera över den uppläggning studien har än att exakt försöka sätta sig in i vilka statistiska tekniker som använts. Med en välgjord design kan ofta pålitliga resultat påräknas även vid mycket enkla statistiska och epidemiologiska analyser.

Slutord

En väl utförd observationsstudie bidrar ofta med väsentlig kunskap som ofta är ganska lätt att tillämpa därför att den är gjord i en reell situation. Det svåra med observationsstudier är att de kräver ganska mycket eftertanke av läsaren.

Referenser

1. Rothman KJ, Greenland S. *Modern Epidemiology*. Philadelphia: Lippincott & Raven; 1998.
2. Elwood JM. *Critical appraisal of epidemiological studies and clinical trials*. Oxford: Oxford University Press; 1997.
3. Greenhalgh T. *How to read a paper. The basics of evidence based medicine*. London: BMJ Books; 2001.
4. Wulff HR, Götzsche PC. *Den kliniska beslutsprocessen*. Lund: Studentlitteratur; 2000.
5. Miettinen OS. The clinical trial as a paradigm for epidemiologic research. *J Clin Epidemiol* 1989;42:491-6.
6. Eriksson M, Hardell L, Berg NO, Moller T, Axelson O. Soft-tissue sarcomas and exposure to chemical substances: a case-referent study. *British Journal of Industrial Medicine* 1981;38:27-33.
7. Nilsson RI, Nordlinder R, Horte LG, Jarvholm B. Leukaemia, lymphoma, and multiple myeloma in seamen on tankers. *Occup Environ Med* 1998;55:517-21.
8. Bhatia R, Lopipero P, Smith AH. Diesel exhaust exposure and lung cancer. *Epidemiology* 1998;9:84-91.
9. Lipsett M, Campleman S. Occupational exposure to diesel exhaust and lung cancer: a meta-analysis. *Am J Public Health* 1999;89:1009-17.
10. Wu AH. Exposure misclassification bias in studies of environmental tobacco smoke and lung cancer. *Environ Health Perspect* 1999;107 Suppl:6873-7.
11. Nyberg F, Agudo A, Boffetta P, Fortes C, Gonzalez CA, Pershagen G. A European validation study of smoking and environmental tobacco smoke exposure in nonsmoking lung cancer cases and controls. *Cancer Causes Control* 1998;9:173-82.
12. Boffetta P, Jarvholm B, Brennan P, Nyrén O. Incidence of lung cancer in a large cohort of non-smoking men from Sweden. *Int J Cancer [serial online]* 2001 Aug [cited 2001 Aug 27]. Available from: URL: <http://www.interscience.wiley.com/cgi-in/abstract/85011323/START>
13. Bansal A, Critchfield GC, Frank TS, Reid JE, Thomas A, Deffenbaugh AM, et al. The predictive value of BRCA1 and BRCA2 mutation testing. *Genetic Testing* 2000;4:45-8.
14. Peto J, Collins N, Barfoot R, Seal S, Warren W, Rahman N, et al.

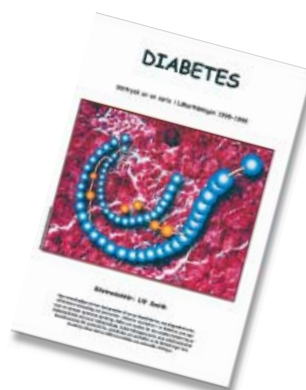
- Prevalence of BRCA1 and BRCA2 gene mutations in patients with early-onset breast cancer. *J Natl Cancer Inst* 1999;91:943-9.
15. Weber BL, Abel KJ, Brody LC, Flejter WL, Chandrasekharappa SC, Couch FJ, et al. Familial breast cancer. Approaching the isolation of a susceptibility gene. *Cancer* 1994;74 Suppl 3:1013-20.
16. Meredith S, Nordman H. Occupational asthma: measures of frequency from four countries. *Thorax* 1996;51:435-40.

Särtryck

Läkartidningen

Nya vetenskapliga rön har lagt grunden till en ny klassifikation, nya diagnoskriterier, effektivare behandling och prevention – inklusive vaccination – av diabetes, som uppvisar en närmast epidemisk spridning. Målen och medlen för den snabba förbättringen av diabetesvården som nu är möjlig belyses i Läkartidningens serie, som också analyserar konsekvenserna för patienterna, sjukvården och samhället av de förändringar som Socialstyrelsen hösten 1999 fastställde som nationella riktlinjer.

Priset är 85 kronor.



Diabetes

Beställer härmed.....ex
av "Diabetes"

.....
namn

.....
adress

.....
postnummer

.....
postadress

**Insändes till Läkartidningen
Box 5603
114 86 Stockholm**

Faxnummer: 08-20 74 35

**www.lakartidningen.se
under särtryck, böcker**

SUMMARY

An observational study requires careful consideration by the reader; merits and weaknesses of observational studies

Bengt Järholm

Läkartidningen 2001;98:5631-7

The observational study has its merits when randomized controlled studies are impossible to perform. This review discusses its merits and weaknesses for a reader of such studies. The reader has to consider the internal and external validity of the study and its role in the establishment of causal relationships and estimations of risks. The reader is recommended to primarily evaluate the quality of study design, confounding and measurement bias, rather than to try to understand if an advanced statistical analysis is correct. The possibilities of misleading results are discussed.

Correspondence: Bengt Järholm, Institutionen för folkhälsa och klinisk medicin, Umeå universitet, SE-901 85 Umeå, Sweden (bengt.jarholm@envmed.umu.se)

Sten Thelander, Statens beredning för medicinsk utvärdering, Stockholm (thelander@sbu.se)

Observationsstudien är central vid flera medicinska frågeställningar

Men det finns en risk för systematiska fel

II Den mångåriga, men ofruktbara, diskussionen om vilken studiedesign som är »bäst« blommade upp i somras, då *New England Journal of Medicine* i samma nummer publicerade två artiklar som inte fann några väsentliga skillnader då behandlingseffekt studerats i randomiserade kontrollerade studier eller i prospektiva kohortstudier [1, 2]. Studierna leddes av en kritisk ledare, skriven av två kända brittiska medicinska metodologer och statistiker [3]. En ledare publicerades också något senare i *British Medical Journal*, i vilken den välgjorda randomiserade studien med rätta framhölls som den mest tillförlitliga [4].

Stora skillnader länder emellan

Sannolikt måste man se diskussionerna i ljuset av de stora skillnader som föreligger mellan USA och övriga industriländers respektive vårdorganisationer. Det amerikanska fragmenterade systemet med starka vinstintressen i alla led, och med en ofta mycket bristfällig vårdkontinuitet, har gjort det allt svårare att genomföra stora randomiserade studier. Federala satsningar på randomiserade undersökningar inom ramen för verksamheten inom National Institutes of Health är omfattande, men otillräckliga, och inte sällan förenade med generaliserbarhetsproblem. Detta gäller speciellt studier av långtidsbehandlingar för kroniska sjukdomar. Läkemedelsstudier är oftast industrifinansierade och har huvudsyftet att ge dokumentation som underlag för kontrollmyndigheternas bedömning av nya preparat.

I ett försök att kompensera för dessa svårigheter har man utvecklat det som brukar kallas »outcome research« [5]. Denna syftar till att utifrån patientadministrativa databaser dra slutsatser om behandlingars effektivitet. Speciellt har man försökt använda de data som samlats in som underlag för att debitera försäkringsbolagen. Dessa data är ju inte alls insamlade i syfte att dokumentera behandlingseffekter. Det är inte slumpen som avgör vilken behandling en viss patient får, utan en mängd mer eller mindre okända faktorer som man sedan försöker korrigera för i avancerade statistiska modeller. Såväl inom som utom USA råder en stark skepsis mot denna typ av studier och de slutsatser som kan dras av dessa [6-8].

Av betydligt större värde är de patientdatabaser som byggts upp i åtskilliga länder, inklusive USA [9]. Dessa initi-

SAMMANFATTAT

Icke-experimentella studier, ofta sammanfattade som observationsstudier, är centrala för många medicinska frågeställningar.

Kvaliteten på observationsstudier har blivit allt bättre.

Hur välgjorda observationsstudier än är kan man inte bortse ifrån risken för systematiska fel.

Observationsstudiernas viktigaste frågeställningar rör riskvärderingar.

Resultaten från observationsstudier bör bekräftas av randomiserade studier.

Randomiserade studier bör ofta kompletteras med observationsstudier för att öka generaliserbarheten av resultaten.

Serie: Evidensbaserad medicin

eras ofta lokalt/regionalt eller av specialistföreningar. I Sverige finner man motsvarande register inom Socialstyrelsens MARS-databas (Medical Access and Result System). Eftersom data oftast rapporteras på klinisknivå är det dock svårt att använda informationen mer än som en jämförelse mellan olika värdenheter. Det är däremot inte möjligt att dra slutsatser om olika behandlingars inbördes effekt och säkerhet.

I Storbritannien finns flera högklassiga register, men när man i en färsk studie försökte använda rutindata från sjukvården i Skottland var informationen genomgående otillräcklig för bestämda slutsatser [10, 11]. Detta innebär att förutsättningen för att patientdataregister skall kunna användas som komplement till randomiserade studier är att de utformas med

II Fakta

Studiedesign

Kontrollerad studie. En viss behandling ges till en grupp patienter, en annan behandling eller ingen behandling alls ges till en annan patientgrupp som man på olika sätt försöker få så likartad den första gruppen som möjligt. Detta kan ske genom matchning för olika egenskaper eller genom fördelning baserad på besöksdag, födelsedag, vårdavdelning eller liknande. Alla dessa sätt att fördela patienter kan potentiellt påverkas av behandlaren, så att grupperna i något viktigt avseende skiljer sig åt.

Randomisering, dvs slumpfördelning till de jämförda grupperna, är det enda sättet att undvika systematiska fel, »bias«, i fördelningen. Randomiseringen gör dock inte med någon automatik grupperna lika, men ju större antal patienter som ingår, desto mindre är risken att slumpen gör grupperna signifikant olika i väsentliga avseenden.

Observationsstudier kallas ibland också epidemiologiska studier, men det mest generella draget hos denna studiekategori är att den inte är experimentell. Huvudkategorierna bland observationsstudier är kohort- respektive fall-kontrollstudier. Båda dessa kan vara antingen prospektiva eller retrospektiva.

Kohortstudier identifierar en population som utsatts för den exponering man vill studera och jämför med en annan kohort som inte exponerats. Sedan följs de båda kohorterna över tid, och utfallet i de olika antalet sjukdomsfall eller dödsfall jämförs.

Fall-kontrollstudien identifierar en grupp individer som utvecklats en sjukdom och som utsatts för en viss exponering. Sedan försöker man matcha dessa med en annan grupp som inte exponerats men i övrigt är så lik den identifierade gruppen som möjligt. Förekomsten av sjukdom hos dessa studeras därefter. Om färre individer i kontrollgruppen utvecklar sjukdom kan detta tala för att en identifierad faktor i fallgruppen medför ökad risk för en viss sjukdom.

stor detaljrikedom, och att avidentifierade data på den enskilda patientnivån blir tillgängliga.

Skepsis mot observationsstudiers värde

Inom den lösligt organiserade rörelse som brukar kallas evidensbaserad medicin (EBM) finns en utbredd skepsis mot observationsstudiernas värde när det gäller att dra slutsatser om behandlingseffekt [12-14]. Problemet med denna kritik är att den ofta drar alla typer av observationsstudier över en kam. I de klassiska studierna från 1970-talet jämfördes de randomiserade studierna med fallserier eller historiska kontroller [15-17]. Då var det föga överraskande att sådana studietyper visade större behandlingseffekter än dem man såg i de randomiserade studierna. Det är ett genomgående resultat från olika forskargrupper att när observationsstudien är utformad som en prospektiv kohortstudie, företrädesvis baserad på de patienter som uppfyllde randomiseringskriterier men som av olika anledningar inte randomiserades, finner man inga systematiska skillnader i effektens riktning eller storlek [1, 2, 18-20].

Begränsningar i den randomiserade studiedesignen

Den experimentella modellen innebär att man försöker hålla allt utom det man önskar studera så standardiserat och enhet-

ligt som möjligt. Detta gör att möjligheterna att upptäcka eventuella skillnader ökar; dock på bekostnad av möjligheterna att generalisera till andra grupper än de som studerats. Man brukar säga att den interna validiteten är hög, men att den externa validiteten kan vara begränsad. Försök har gjorts att skapa modeller för att integrera information från randomiserade studier och observationsstudier i så kallad »cross-design synthesis«. Dessa teoretiskt tilltalande försök har inte gett några praktiskt användbara tillämpningar.

När väl en effekt påvisats i en randomiserad studie, eller i en metaanalys av flera studier, kan även effekten i mer selekterade populationer undersökas med samma metod. Dessa stora, enkla, randomiserade studier har med stor framgång tillämpats inom bl a onkologi och kardiologi. Flera tiotusentals patienter har randomiserats till två eller tre behandlingsgrupper. I dessa grupper har patienter med mycket varierande egenskaper inkluderats. Förutom den studerade behandlingen har övriga behandlingar givits enligt respektive kliniks egna rutiner. Med så stora grupper, och förutsatt att prognostiskt intressanta subgrupper definierats före studien, har undersökningar av skillnader mellan dessa subgruppers behandlingsresultat kunnat göras.

Finansieringsproblem, logistiska svårigheter och trögheter i forskningsorganisation har dock gjort att de flesta behandlingar i sjukvården inte undersökts i sådana gigantiska studier.

För den behandlande klinikern är därför en återkommande fråga i vilken utsträckning som kliniska behandlingsresultat är tillämpliga för en viss patient. Ålder, kön och annan samtidig sjukdom, liksom behandlingen av denna, är faktorer som kan vara annorlunda än de som kännetecknade deltagarna i de randomiserade undersökningarna. Även om dessa faktorer är desamma kan den aktuella sjukdomen vara lindrigare eller allvarligare än hos de studerade patienterna. En behandling kanske är mindre verksam vid långt gånget sjukdom. Eller, tvärtom, kan sjukdomen vara alltför beskedlig för att behandlingen skall kunna ge någon specifik effekt [21]. Allvarliga men ovanliga komplikationer av behandlingen kanske inte längre är acceptabla om den form av sjukdomen som den aktuella patienten har är förhållandevis godartad.

Välgjorda observationsstudier kan ge värdefull kompletterande information om behandlingseffekter i mindre selekterade patientpopulationer än de som studeras i randomiserade prövningar. Väsentlig kritik har framförts mot vissa uttorkningar av paradigmat evidensbaserad medicin [22, 23]. Denna kritik gäller dock bara de, lyckligtvis, få som mer eller mindre öppet förnekar behovet av andra studietyper än den randomiserade, kontrollerade prövningen.

Observationsstudiernas viktigaste användningsområden

De huvudsakliga frågeställningarna vid observationsstudier är förekomsten av riskfaktorer antingen i miljö eller som följd av en oavsiktlig eller avsiktlig exponering för en faktor som kan utgöra en risk [24]. Upptäckten av allvarliga men ovanliga biverkningar av läkemedel eller toxiner kan i princip ske enbart med hjälp av observationsdata. Dessa kan vara i form av obligatorisk eller spontan rapportering av oönskade reaktioner på behandling, där sedan farmakoepidemiologer söker efter ett mönster. Av särskild vikt är symtom som inte hör till läkemedlets farmakologiska effekter och som inte brukar ses vid den sjukdom som behandlas.

Identifieringen av riskfaktorer i omgivningen är svårare. Små riskökningar kan lika väl bero på okända bakomliggande faktorer, »confounders«, som på en faktisk effekt. Reproducerade riskökningar i andra populationer kan öka trovärdigheten, men om samma studiemetod tillämpas kan samma »confounder« vara verksam i alla. Ett dos-effekts samband ger

ett starkt stöd för att ett orsakssamband föreligger. Likaså är stora riskökningar ofta svårare att förklara med »confounders«.

Det finns ingen enighet om vad som skall menas med små respektive stora riskökningar. Vissa brukar ange en oddsratio på över 2 eller 3 för att man skall kunna fästa någon större vikt vid en observationsstudie [14]. Andra avvisar denna typ av gränssättning och vill i stället väga ihop resultatet med patofysiologisk rimlighet och liknande [22, 23]. Det finns inte heller någon enighet om värdet av att göra metaanalyser av observationsstudier [25-27]. Antalet publicerade metaanalyser av observationsstudier ökar dock mycket snabbt. En stor försiktighet är dock att rekommendera vad gäller tolkningen av sådana metaanalyser, kanske speciellt baserade på fall-kontrollstudier [27]. I flera jämförande studier har man funnit betydligt större effektstorlekar i fall-kontrollstudier än i kohortstudier med samma frågeställning.

Utmaning

Det är i allmänhet betydligt svårare att bedöma observationsstudier än randomiserade behandlingsstudier. Inte sällan används sofistikerade statistiska metoder för att försöka korrigera för »confounders«. Sådana metoder ingår sällan i den kliniska forskarutbildningen, och aldrig i den medicinska grundutbildningen. Om man kan vara tveksam till hurtiga påståenden om att evidensbaserad medicin kan läras in på någon dags kurs vore motsvarande påstående om klassisk epidemiologi enbart löjeväckande. Det finns däremot ett stort behov av ökade kunskaper och färdigheter i utformningen speciellt av kohortstudier inom den kliniska forskarvärlden. Det förefaller helt klart att många viktiga kliniska frågeställningar inte kan besvaras enbart med hjälp av den randomiserade prövningen [28-30].

Det är en stor utmaning att tillämpa eller utveckla metoder för att fördjupa och bredda den information om behandlingseffekter som ursprungligen påvisats med hjälp av den randomiserade studieuppläggnen.

Referenser

1. Benson K, Hartz A. A comparison of observational studies and randomized controlled trials. *N Engl J Med* 2000; 342: 1878-86.
2. Concato J, Shav N, Horwitz R. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med* 2000; 342: 1887-92.
3. Pocock S, Elbourne D. Randomized trials or observational tribulations? *N Engl J Med* 2000; 342: 1907-9.
4. Barton S. Which clinical studies provide the best evidence? The best RCT still trumps the best observational study. *BMJ* 2000; 321: 255-6.
6. Byar D. Problems with using observational databases to compare treatments. *Stat Med* 1991; 10: 663-6.
8. Sheldon T. Please bypass the PORT. *BMJ* 1994; 309: 142-3.
9. Black N. High-quality clinical databases: breaking down barriers. *Lancet* 1999; 353: 1205-6.
10. Lewsey J, Murray G, Leyland A, Boddy F. Comparing outcomes of percutaneous transluminal coronary angioplasty with coronary artery bypass grafting – can routine health service data complement and enhance randomized controlled trials? *Eur Heart J* 1999; 20: 1731-5.
12. Kunz R, Oxman A. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. *BMJ* 1998; 317: 1185-90.
14. Sackett D, Struss S, Richardson S, Rosenberg W, Haynes B. Evidence Based Medicine - How to practice and teach EBM. 2nd ed. London: Churchill Livingstone, 2000.
19. McKee M, Britton A, Black N, McPherson K, Sanderson C, Bain C. Interpreting the evidence: choosing between randomised and non-randomised studies. *BMJ* 1999; 319: 312-5.
20. MacLehose R, Reeves B, Harvey I, Sheldon T, Russell I, Black A. A systematic review of comparisons of effect sizes derived from ran-

domised and non-randomised studies. *Health Technol Assess* 2000; 4: 1-168.

21. Ioannides J, Lau J. The impact of high-risk patients on the results of clinical trials. *J Clin Epidemiol* 1997; 50: 1089-98.
22. Feinstein A, Horwitz R. Problems in the »evidence« of »evidence-based medicine«. *Am J Med* 1997; 103: 529-35.
23. Vandenbroucke J. Observational research and evidence-based medicine: what should we teach young physicians? *J Clin Epidemiol* 1998; 51: 467-72.
26. Shapiro S. Is meta-analysis a valid approach to the evaluation of small effects in observational studies. *J Clin Epidemiol* 1997; 50: 223-9.
27. Egger M, Schneider M, Smith G. Spurious precision? Meta-analysis of observational studies. *BMJ* 1998; 1998: 140-4.
29. Naylor D. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet* 1995; 345: 840-2.
30. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996; 312: 1215-8.

Fullständig referenslista kan erhållas från författaren.

Lars Werkö, professor, Stockholm

Väl underbyggda riktlinjer måste vara färskvara

Men kontinuerlig uppdatering befriar inte läkaren från att tänka själv

II I kölvattnet till utvecklingen av evidensbaserad medicin har behovet av mer eller mindre officiella riktlinjer, »guidelines«, blivit uppenbart. Det första försöket utgjordes av konsensusdokument, som emellertid fått mindre betydelse på grund av deras begränsade vetenskapliga förankring. I fortsättningen har såväl i Sverige som internationellt olika organisationer eller myndigheter ägnat sig åt att producera bättre underbyggda nationella eller regionala riktlinjer, många gånger avsedda som underlag för vårdprogram.

Detta har inte alltid mottagits med entusiasm. I England har ett flertal specialister samlat sig till en samlad kritik mot den kontroll av verksamheten som man anser riktlinjer vara [1]. I synnerhet kardiologerna anser sig själva kunna bedöma den litteratur som dessa riktlinjer baseras på utan pekfinger från en myndighet som man antar har som huvudsakligt mål att spara pengar för National Health Service [2].

Det är dessutom alltmer uppenbart att svårigheter föreligger att hålla såväl centralt framtagna riktlinjer som systematiska översikter eller andra sammanställningar aktuella (se faktaruta).

En viss fördröjning måste accepteras

Det är emellertid omöjligt för den enskilde läkaren att själv följa alla nya betydelsefulla rön, vilket är nödvändigt om kvaliteten i sjukvården skall kunna vidmakthållas. Detta gäller i synnerhet som det är fråga om att inte endast läsa och tillämpa utan också att kritiskt granska och värdera de framsteg som publiceras. I Sverige har en viktig uppgift för Statens beredning för medicinsk utvärdering (SBU) varit att sammanställa det vetenskapliga underlaget i en rapport på ett sådant sätt att den kan användas som underlag för utarbetandet av lokala eller nationella vårdprogram. Tyvärr innebär detta en fördröjning av tillämpningen av viktiga nyheter eftersom redan utarbetandet av en rapport tar (alltför) lång tid. Ett sätt att minska denna fördröjning har varit det »snabbrapportssystem« som SBU Alert utgör, men även denna verksamhet tar tid.

Det torde därför vara lika bra att från början medge att sjukvården måste acceptera den fördröjning som bearbetande av allt nytt material innebär. Den gamla regeln att använda metoder man är väl förtrogen med, hellre än att snabbt försöka tillämpa allt nytt, gäller fortfarande – det nya kan vara

Evidensbaserad medicin

riskabelt att införa innan det är tillräckligt utvärderat. Att tillämpa kliniska riktlinjer eller vårdprogram är emellertid inte heller utan sina problem. Alla procedurer i vårdprogramarbetet kräver tid, vilket kan innebära betydande fördröjning. Inom områden med snabb utveckling av nya metoder medför detta att även ett färskt dokument redan när det publiceras kan vara föråldrat och i behov av omarbetning. Det är dock bättre i rutinsjukvården att använda säkra metoder än att arbeta i den absoluta nyhetsfronten.

Rutiner för uppdatering

Ytterligare ett problem är att varje systematisk översikt måste följas och uppdateras. SBU har sedan länge infört som rutin att varje projektgrupp ansvarig för en rapport kallas samman två till fem år efter det att rapporten publicerats för att bedöma behovet av uppdatering. Efter en snabb sökning av litteraturen kan man sedan avgöra behovet av revision.

På samma sätt har den amerikanska US Agency for Healthcare Research and Quality (AHRQ) följt upp sina rekommendationer [3]. Shekelle och medarbetare utarbetade kriterier för när en rekommendation behövde uppdateras. I samarbete med de 170 experter som från början utarbetat de 17 guidelines som AHRQ publicerat granskades den nyutkomna litteraturen fram till mars 2000. Man fick 6 994 träffar i databaserna, som efter granskning resulterade i 173 artiklar plus 159 nya rekommendationer inom motsvarande ämnesområde.

Nästan tre fjärdedelar av experterna i de ursprungliga projektgrupperna deltog i denna nya utvärdering. De fann att sju guidelines fordrade en större revidering på grund av ny kunskap, sex behövde mindre förändringar, medan tre ansågs fortfarande vara i överensstämmelse med det aktuella kunskapsläget. I ett fall kunde man inte komma överens om huruvida det behövdes någon omarbetning.

En matematisk analys av livslängden för organisationens rekommendationer gav vid handen att de inte längre var till-

II Fakta

Många olika typer av riktlinjer

Systematiska översikter

Sammanställning av aktuell vetenskaplig litteratur som funnits i angivna databaser.

Kliniska riktlinjer

Rekommendationer av centrala eller lokala myndigheter eller professionella sällskap, oftast – men inte alltid – baserade på systematiska översikter.

Vårdprogram

Anvisningar om diagnos och behandling av ett visst sjukdomstillstånd, lokalt eller centralt framtagna och oftast baserade på systematiska översikter eller kliniska riktlinjer.

Konsensusdokument

Anvisningar och rekommendationer framtagna i diskussioner mellan inkallade experter med begränsad tillgång till systematiska översikter och baserade på överensstämmande åsikter hos experterna.

»Guidelines«

Engelska/amerikanska kliniska riktlinjer eller vårdprogram.

förlitliga efter i medeltal 5,8 år (5,0–6,6 år). Den tid som 90 procent av rekommendationerna fortfarande var användbara var 3,6 år (2,6–4,6 år). Denna beräkning ledde till rekommendationen att riktlinjer borde granskas för eventuell uppdatering vart tredje år.

Utän att betvivla detta resultat måste man fråga sig om det verkligen är möjligt att både försöka täcka hela det spektrum som sjukvården idag omfattar med aktuella riktlinjer och också revidera dessa vart tredje år. Leder inte detta till en enorm byråkratisk organisation som hotar att kväva hela verksamheten? I en ledarkommentar utvecklar George P Browman de svårigheter som föreligger när man fordrar systematiska sammanställningar som bakgrund till alla råd i de riktlinjer som har blivit så populära. Han anser att dessa svårigheter är så stora att de rekommendationer som Shekelle och medarbetare lagt fram snarare skall ses som ett intressant inlägg i den metodologiska debatten än som något som bör följas i det praktiska arbetet [4].

Det finns ju dessutom andra sätt att hålla översikter à jour med utvecklingen. Det sätt på vilket Cochrane Collaboration publicerar sina systematiska översikter på Internet, med uppdatering varje kvartal, innebär en mindre börda för dem som ansvarar för utarbetandet av rekommendationer, samtidigt som det ställer större anspråk på användarna, som måste följa utvecklingen kontinuerligt.

Inte heller Cochrane är perfekt

Svårigheterna illustreras av den granskning som en grupp av Cochrane-medarbetare gjort av ett urval av de systematiska översikter som publicerades 1998 [5]. I debatten om kritiska sammanfattningar av aktuellt vetande har man betraktat de riktlinjer som gjorts inom Cochrane Collaboration med dess rigorösa bestämmelser som den gyllene standard som alla

skulle vilja leva upp till. Men, som Olsen och medarbetare anger [5], ingenting är så bra att det inte kan göras bättre. I de 53 översikter som publicerades 1998 visade det sig att inte mindre än 15 – nära en tredjedel – uppvisade vad gruppen kallade större, »major«, problem. Det vetenskapliga underlaget stödde inte de slutsatser som dragits i nio sammanställningar, medan sättet att rapportera lämnade en del övrigt att önska i tolv rapporter. Gruppen påpekar att organisationen utvecklats sedan 1998, och att ett kontinuerligt arbete med kvalitetskontroll för att minimera subjektiva inflytanden vid utvärdering av litteraturen också förbättrar resultatet kontinuerligt.

Gemensamt för samtliga rapporter där man fann tveksamheter var att slutsatserna var mer gynnsamma än vad underlaget tillät. Detta är särskilt anmärkningsvärt, eftersom det är just i detta avseende som kritik alltid riktas mot de diskussioner, resultat och slutsatser som läggs fram av personer knutna till företag eller som på annat sätt är personligt intresserade av positiva resultat. Det verkar som om läkare när de sätts att granska underlag till åtgärder som skall gagna deras patienter – även när de agerar långt ifrån den direkta patientkontakten – tenderar att vara mer optimistiska och mer positiva än som strikt objektivitet motiverar. Klinisk träning kanske medför en omedveten, men nödvändig, positiv inställning till den verksamhet som skall komma en lidande medmänniska till godo. Detta gör det än mer nödvändigt att hårt granska vetenskapliga resultat – optimismen får uttryckas vid sjukbädden, efter det att den kalla vetenskapliga bedömningen har sagt sitt.

Cochranes mammografirapport

Det är av intresse att i den granskningsgrupp som kritiserar de tidigare Cochrane-publicationerna ingår Ole Olsen och Peter Götzsche, som bägge står för den Cochrane-rapport om mammografi som nu utsätts för diskussion och kritik [6, 7]. Kritiken har varit särskilt inriktad på hur man bedömt flera av de studier som tidigare ansetts som föredömliga. Olsen och Götzsche har beskyllts för att ha tagit alltför stor hänsyn till »teknikaliteter« i sitt underkännande av vissa studier. Kritiken har också stötts av det faktum att Cochrane Breast Cancer Group, som ansvarar för den redaktionella bearbetningen av bröstcancer rapporter, inte följt Olsens och Götzsches skrivning utan själva lagt till några uttalanden som mildrar eller eliminerar kritiken.

Däremot har redaktören för Lancet i en underledarkommentar understött Olsen och Götzsche [8], vilket har särskilt förtörnat Alessandro Liberati, ordförande i Cochrane Breast Cancer Group. Denne har uttalat sig dels i BMJ [9], dels i ett brev till Lancet, där han framhåller:

»But we have substantial disagreement on the interpretation of the data on the impact of mammography screening on surgery and treatment so we have offered to wait for further peer review. In order not to delay the Cochrane review, we suggested that it should go ahead and be updated in subsequent issues of the Cochrane library after further peer review.«

Liberati säger också: »Even after 30 years we don't really have the data we need – particularly for many outcomes that are important to screening women – in the case of mammography screening.«

Många av dem som stöder mammografirörelsen har framför allt kritiserat den alltför stora betydelsen av det som betecknas »teknikaliteter« i danskarnas kritiska granskning. Skillnaden mellan en systematisk, kritisk sammanställning och ett referat av ursprungsförfattarnas slutsatser är emellertid just en kontroll och bedömning av »teknikaliteter«. De tillägg bröstcancergruppen har gjort har gjorts i avvaktan på ytterligare »peer review« av vissa av de arbeten som Olsen och

Götzsche inkluderat i sitt negativa uttalande om mammografierna.

Kvalitetsarbetet är viktigt

En undersökning som inte följer det protokoll man själv varit med om att författa, även om det endast är i vissa detaljer, kan inte anses vara trovärdig, hur politiskt korrekt resultatet än blir. Kvalitetsarbete inom de viktiga organisationer som förser sjukvården med sammanställningar av vetenskapliga underlag för beslut måste ske kontinuerligt och intensivt för att bibehålla sin trovärdighet.

Det är därför uppmuntrande att Cochrane Collaboration, trots den snabbt växande bördan, särskilt betonar kvalitetsarbetet. Det finns för närvarande mer än 1 000 publicerade sammanställningar och 800 protokoll för pågående granskningar inom organisationen. Det torde ändå finnas behov av minst lika många till, utöver det hela tiden brännande kravet att hålla hittillsvarande rapporter aktuella. Som Clarke och Langhorne [6] anger: »Achieving this level of coverage while keeping all reviews up to date is perhaps the biggest challenge in the years to come.«

Det finns en rädsla att bryta mot rekommendationer

Det är alltså svårt att åstadkomma heltäckande kliniska riktlinjer och se till att de verkligen hålls aktuella. Hur viktiga är då dessa riktlinjer?

Många läkare är rädda för att förekomsten av rekommendationer eller vårdprogram kan leda till att de blir anmälda för brister i sin yrkesutövning om de inte följer de senaste råden i sin verksamhet [1, 10]. I USA är detta ännu mera aktuellt än i Sverige. Den som är ansvarig för arbetet med guidelines i den amerikanska organisationen för internmedicin, American College of Physicians – American Society of Internal Medicine, Vincenza Snow, har gjort en enkätstudie bland invärtesmedicinare i USA som skötte patienter med diabetes, omfattande 85 läkare och 1 755 patientmöten [10]. Hon undersökte hur troget man följt riktlinjer för »best practices«. Det visade sig att detta långt ifrån alltid var fallet. Man var dock angelägen om att följa råden i guidelines, dock med de modifieringar som den aktuella patientsituationen kunde motivera.

Många avvikelser förekom således i praktiken, men de kunde motiveras med den kliniska bild som den aktuella patienten uppvisade. Avsteg från riktlinjer behöver alltså inte betyda att man inte följer dem, snarare visar de att det är en självständigt tänkande kliniker som står för vården.

Det sammanlagda intrycket av diskussionen om nytta eller svårigheter med systematiska översikter eller kliniska riktlinjer blir att de kan vara till hjälp, men att de inte befriar läkaren från att tänka själv. Svårigheter med att täcka hela det medicinska spektrum som en läkare möter leder till att endast vissa stora områden kan omfattas. För bästa möjliga resultat måste systematiska sammanställningar hållas aktuella, vilket bäst torde kunna ske genom elektronisk publicering i kombination med kontinuerlig revision. Ett sådant hanterande av problemet skulle också kunna hålla otillbörligt inflytande från företag och »missionärer« till ett minimum.

Referenser

1. Miles A, Hampton JR, Hurwitz B, editors. NICE, CHI and the NHS Reforms. London: Aesculapi Medical Press; 2000.
2. Hampton JR. Evidence, guidelines, audit and cardiology: principles and problems in secondary and tertiary care. In: Miles A, Hampton JR, Hurwitz B, editors. NICE, CHI and the NHS Reforms. London: Aesculapi Medical Press; 2000. p. 79-88.
3. Shekelle PG, Ortiz E, Rhodes S, Morton SC, Eccles MP, Grimshaw JM, et al. Validity of the agency for healthcare research and quality

clinical practice guidelines. How quickly do guidelines become outdated? JAMA 2001;286:1461-7.

4. Brownman GP. Development and aftercare of clinical guidelines. The balance between rigor and pragmatism. JAMA 2001;286:1509-11.
5. Olsen O, Middleton P, Ezzo J, Götzsche PC, Hadhazy V, Herxheimer A, et al. Quality of Cochrane reviews: assessment of sample from 1998. BMJ 2001;323:829-32.
6. Clarke M, Langhorne P. Revisiting the Cochrane Collaboration. Meeting the challenge of Archie Cochrane – and facing up to some new ones. BMJ 2001;323:821.
7. Olsen O, Götzsche PC. Cochrane review on screening for breast cancer with mammography. Lancet 2001;358:1340-2.
8. Horton R. Screening mammography – an overview revisited. Lancet 2001;358:1340-2.
9. Mayor S. Row over breast cancer screening shows that scientists bring »some subjectivity into their work«. BMJ 2001;323:956.
10. Larkin M. Noncompliance with »best practices« may be justifiable. Lancet 2001;358:1433.

»Medicinsk kommentar« är Läkartidningens forum för signerade medicinska ledare. Merparten av dessa är beställda av redaktionen, och vi välkomnar förslag om aktuella frågor som bör tas upp i denna form. Vi vill där även fånga in och belysa aktuella medicinska rön presenterade annorstädes.

Finns något i din specialitet att kommentera? Ta kontakt med redaktionen innan du börjar skriva för att undvika dubbelarbete!

Evidensbaserad medicin – ny kurs på läkarutbildningen

Kurser i evidensbaserad medicin (EBM) planeras vid flera av landets läkarutbildningar, och i Umeå har kurser redan startats. Huvudsyftet är att introducera EBM:s grundprinciper, men framför allt att tidigt träna förmågan till att självständigt söka kunskap och att kritiskt granska faktaunderlaget för det kliniska handlandet. Utbildningen sker till stora delar som grupparbeten, men också kortare katedrala moment, patientdemonstrationer och rollspel ingår. Ur lärarperspektiv innebär EBM-undervisningen en nytändning, och studenternas prestationer är förbluffande goda. Studenternas synpunkter på kursen redovisas i en separat artikel.

KJELL ASPLUND
professor, medicinska kliniken

ANNA ENGSTRÖM-LAURENT
docent, universitetslektor, medicinska kliniken

ELLINOR ÄDELROTH
docent, universitetslektor, lungkliniken;
samtliga vid Norrlands Universitetssjukhus, Umeå

II I de flesta svenska läkarutbildningar ingår numera ett element av vetenskaplig träning. Utformningen varierar mellan lärosätena. I Umeås nya studieplan kallas momentet *Vetenskaplig baskurs*. Den vetenskapliga baskursen återkommer under hela utbildningen och innefattar eget arbete och valbara fördjupningskurser. Den första terminens första månad ges studenterna en vetenskaplig introduktionskurs som inleder läkarstudierna.

I introduktionskursen *Vetenskapen, människan och samhället* ingår en vecka om evidensbaserad medicin (EBM), något som Umeå är först i landet med. Liknande kurser i EBM planeras nu vid flera studieorter. Den första av våra kurser gick i januari–februari år 2001.

Kritiskt granskande

Kursens övergripande målsättning stämmer överens med definitionen för evidensbaserad medicin: att självständigt och kritiskt kunna granska faktaunderlaget för kliniskt handlande, till gagn för enskilda patienter.

Några delmål är:

- att kunna grundläggande element i evidensbaserad medicin (definitioner, kliniska utfall av betydelse för enskilda patienter, absoluta kontra

relativa risker, klinisk och statistisk signifikans, olika studiers tyngd i det kliniska beslutsfattandet, metaanalyser och kostnadseffektivitet).

- att vara orienterad om syntesmetoder och styrdokument i sjukvården (konventionella kontra systematiska kunskapsöversikter, konsensusdokument, state-of-the-art, nationella kliniska riktlinjer och vårdprogram).
- att söka tillämpa EBM-principerna på exempel ur vardagssjukvården.
- att känna till begränsningar och etiska aspekter inom evidensbaserad medicin i relation till andra komponenter i läkekonsten, samt förhållningssätt när det vetenskapliga underlaget är svagt.
- att vara orienterad om evidensbaserad medicin som bas för prioriteringar.

Ambitionerna för en veckas kurs är stora, men de studenter som väntade sig att läkarstudierna skulle börja med räplugg i anatomi blir lätt förvirrade. Tanken är nämligen att redan den första månaden anslå tonen för hela den kommande läkarutbildningen; det handlar om självständigt kritiskt tänkande och om praktisk tillämpning av EBM-principerna på vanliga kliniska problem.

Pedagogisk undervisning

Före veckan med evidensbaserad medicin börjar har studenterna fått ingående kunskaper i hur man söker medicinsk litteratur. De har också fått orienterande undervisning om medicinsk vetenskapsteori och historia, statistik, forskningsmetodik samt forskningsetik.

Undervisningen är studentaktiverande. För att få liv i utbildningen tillämpas flera pedagogiska metoder:

Se även artikeln på sidan 2179 i detta nummer.

- modifierad fallmetodik (korta fall, inget bakgrundsmaterial) med arbete i grupper om åtta studenter och erfarna och EBM-kunniga kliniker som grupphandledare,
- ett litet antal katedrala presentationer,
- patientdemonstrationer,
- rollspel.

Vi har haft fyra teman under veckan: diagnostik, behandling, hälsoekonomi och syntes (se studenternas artikel). Inom ett avsnitt, hälsoekonomin, har vi varit överambitiösa. Som sjukvården fungerar idag kan det nog vara bra att tidigt introducera studenterna i hälsoekonomiskt tänkande, men våra erfarenheter är att detta måste göras på ett enkelt sätt, nära knutet till vardagssjukvården för att upplevas som relevant.

Patientdemonstrationerna har haft som huvudsyfte att visa hur det går till när patient och läkare tillsammans angriper och löser kliniska problem utifrån EBM-principerna. De har också illustrerat hur komplex den kliniska verkligheten är. Långt ifrån alla problem handlar om att tillämpa evidensbaserad medicin – läkekonsten har också andra viktiga dimensioner.

Rutorna »Diagnostik av aortaaneurysm« och »Medicinsk kontrovers: bröstcancerundersökning« ger två exempel på de uppgifter studenterna ska försöka lösa. Den förstnämnda är en vanlig gruppuppgift och den andra en uppgift kring en medicinsk kontrovers som ska redovisas som rollspel.

Bra studentprestationer

Under EBM-veckan möter studenterna många exempel på hur bräckligt det vetenskapliga underlaget för vårt kliniska handlande kan vara och hur kontroversiella många insatser är. Denna korta kurs handlar ju inte bara om att förmedla fakta. Den vill framför allt skapa en attityd; akademiska studier innebär ett aktivt sökande efter kunskap och ett kritiskt ifrågasättande.

Som lärare med lång erfarenhet av traditionell läkarutbildning är det spän-

Diagnostik av aortaaneurysm

Bukaortaaneurysm (vidgning av bukaorta) kan vara dödligt. Ibland kan patienten söka för buksmärter, men det är också vanligt att ett aneurysm upptäcks när buken palperas vid rutinundersökning. När väl diagnosen misstänks brukar man göra ultraljudsundersökning för att bekräfta den. Det finns dock en rad svårigheter med den kliniska diagnostiken som vi ber att du reder ut:

- När misstanken uppkommer palperas buken ofta flera gånger och av flera doktorer för att bedömningen ska bli säkrare. Är detta riskfritt eller innebär en bukpalpation att risken ökar för att aneurysmet brister?
- Många äldre har en slingrig, vidgad aorta och det kan vara svårt att avgöra vad som är normalt för åldern och vad som är en sjuklig vidgning av aorta – ett riktigt aneurysm. Hur ska man avgränsa diagnosen aortaaneurysm?
- Vilken precision har bukpalpation för diagnostisera aortaaneurysm? Vet man något om sensitivitet, specificitet och positivt prediktivt värde?

Medicinsk kontrovers: bröstcancerundersökning

Detta fall rör en medicinsk kontrovers där fakta måste blandas med värderingar. Det finns inte ett enda »rätt svar« på frågorna.

Penetrera först den vetenskapliga litteraturen och inventera olika argument. Redovisa sedan i form av ett rollspel. Välj en i gruppen som står för en åsikt i den huvudfråga som ställs och en som står för motsatt åsikt. Vi ordnar en kort paneldebatt i varje fråga.

En 46-årig kvinnlig lärare utan ärftlighet för bröstcancer ber om råd. Hon har kallats till en röntgenundersökning av bröstet för att tidigt upptäcka bröstcancer. Hon har dock läst att detta är en kontroversiell undersökning, särskilt i den ålder hon befinner sig. Hon har med sig en serie artiklar ur Svenska Dagbladet från sommaren 2000, där det vetenskapliga underlaget för dessa screeningundersökningar ifrågasätts. Hon berättar dessutom om en kollega på jobbet som fått besked om att röntgenundersökningen var helt normal och som ett halvår senare själv upptäckte en ganska stor bröstcancer.

Nu frågar hon: varför är man så oense om värdet av denna undersökning? Finns det risker med undersökningen? (I den kallelse hon fått nämns ingenting om eventuella risker.) Hur bra är metoden egentligen för att upptäcka bröstcancer?

För rollspelet: Vilka är argumenten för respektive emot bröstcancerundersökningar hos kvinnor under 50 år? Ta med medicinska, ekonomiska och etiska/sociala aspekter i resonemanget.

Två exempel på uppgifter som studenterna ska lösa på kursen i evidensbaserad medicin på läkarutbildningen.

ande att ge sig in på ett nytt lärostoff och göra det i nya undervisningsformer. Vi förväns över vad studenterna klarar av redan under sin första studiemånad vid universitet. De är självständiga, påhittiga och imponerande receptiva.

Vi hoppas att undervisningen i evidensbaserad medicin på grundnivå kommer att resultera i läkare med goda kunskaper och med stor förmåga att med ett öppet och kritiskt sinne ta till sig nytt vetande. •

EBM-kurs gav kritiskt tänkande

Efter bara tre veckor på läkarutbildningen i Umeå får studenterna gå en kurs i evidensbaserad medicin. Här presenterar några studenter sina upplevelser av kursen.

SARA WIRD

ANNELIE MELANDER

KAROLINA BOMAN

ANDERS RUUTH

DANIEL FRÖDING

samtliga studenter vid läkarprogrammet, Umeå universitet

■ En septembermåndag 2001 sammanfördes sju studenter för första gången. Uppgiften vi ställdes inför var att via databaser ta reda på bästa tänkbara metod för att diagnostisera hjärtinfarkt. Dagarna därefter behandlade vi lungcancerdiagnostik, hälseneruptur, bröstcancer-screening och familjär amyloidpolyneuropati. Tempot var högt eftersom mycket skulle hinnas med på en vecka. En ortoped vi pratade med angående hälsenerupturen påpekade vårt problem/svårighet/utmaning med orden: Ni vet ju ingenting! Och han hade rätt. Han kunde inte veta att vi gått enbart tre veckor på läkarprogrammet. De tidigare erfarenheter vi hade av arbete inom vården var högst varierande. Vad gäller vår specifika grupp var en person undersköterska sedan tidigare och några hade jobbat extra inom vård och omsorg.

Under de första stapplande veckorna hängde EBM-begreppet över oss konstant. Vad det egentligen innebar var det nog ingen som greppade. Ett sätt att förbereda oss för EBM-metodiken var att under en vecka lära oss söka och värdera information ur en mängd databaser, tidskrifter och andra källor. Utan detta moment hade den efterföljande EBM-veckan varit mer påfrestande och frustrerande än lärorik.

Föreläsare nykära i EBM

Upplägget på EBM-veckan hade ett ambitiöst och fullspäckat schema. Morgnarna inleddes med gruppredovisningar av den föregående dagens fall. Därefter följde en föreläsning som förberedelse för eftermiddagens grupparbete. Lunchen spenderades till att lära känna de nya klasskompisarna i kön till mikrovågsugnen. De olika momenten hade alla sina poänger och kompletterade varandra bra.

Föreläsningarna var av skiftande karaktär och kvalitet. De var långa men generellt sett nödvändiga. Det som var positivt och som återfanns i praktiskt taget

alla föreläsningar, var kontakten med det kommande arbetslivet som läkare. Måndag till onsdag gick vi igenom diagnostik, behandling och hälsoekonomi.

Torsdagens föreläsning och grupparbete fokuserade syntes av de tre ovan nämnda momenten med en kritisk granskning av EBM. Det märktes att föreläsarna var mycket entusiastiska och näst intill nykära i EBM. Föreläsningarna kryddades och illustrerades med patientbesök där föreläsaren i samtal med patient visade hur en konsultation kan gå till. Det var uppskattade inslag och bra för alla, och speciellt givande för dem av oss som aldrig befunnit sig i en konsultationssituation som patient tidigare.

Fallredovisningarna genomfördes på olika sätt men upplevdes nog av alla som både spännande och skrämmande. Att inför 80 nya klasskamrater prata om ett ämne man dagen innan inte vetat något om var stimulerande. Fallen redovisades bland annat med OH-visningar, paneldebatt och små sketcher eller rollspel.

Mycket faktaletande

Inför grupparbetena fanns en viss oro i gruppen eftersom många var vana att sedan tidigare ta på sig ledaransvar. EBM-veckan blev därför även en träning i grupparbete, och detta var minst lika värdefullt som kunskaperna i EBM.

Grupparbetet gick dessutom bra. Vi lärde oss att fördela arbetet och använde en stor del av eftermiddagarna att söka nödvändiga fakta, som vi sedan sammanfattade inför hela klassen. Utöver de sökmetoder vi hade lärt oss tidigare pratade vi med läkare och ringde labb. Vi fick bra hjälp från våra handledare i alla våra funderingar. Den tryckta litteraturen vi använde under veckan var »Evi-

Se även artikeln på sidan 2176 i detta nummer.

dence-based medicine« av Sackett m fl. Våra kursansvariga propagerade starkt för denna blå EBM-bibel. Vi upplevde den inte alls lika positiv och för oss relevant. De kompendier vi fick ut i samband med föreläsningarna under veckan räckte gott och väl för oss. Vi kan dock tänka oss att den lilla blå boken kommer till användning i framtiden.

Allt detta mynnade ut i en skriftlig tentamen och en examinationsuppgift. Examinationsuppgiften hade samma upplägg som veckans grupparbeten men var mer omfattande. Den redovisades dessutom både skriftligt och muntligt.

Avslutningsvis kan vi säga att kursen kändes väldigt relevant i tiden, trots våra ringa medicinska förkunskaper. Vi kände att vi läste på läkarlinjen även om det inte var lika konkreta kunskaper som vi skulle fått av en anatomikers. Trots att vi upplevde kursen positiv så antar vi att våra utvärderingar leder till ständig utveckling av kursens innehåll.

Nu, mitt ibland mitokondrier, bakterier och DNA under pågående cellkurs, kan EBM-metodiken kännas avlägsen. Men i vår vardag använder vi oss av kritiskt tänkande i större utsträckning än förut. I vårt informationsökande i och läsande av vetenskapliga artiklar finns det kritiska förhållningssättet. Vi har lärt oss att artiklarna, lika lite som våra läroböcker och vetenskapliga rön som publiceras i massmedierna, inte alltid är sanningen eller kanske ens redovisar de senaste evidensbaserade forskningsresultaten. Hur den där septemberveckan med hjärtinfarktdiagnostik, Skelleftesjuka och kontrollerade kliniska studier kommer att påverka vår fortsatta utbildning och framtida yrkesroll återstår att se. •

Redovisning av examensuppgift i form av rollspel där artikelförfattarna deltar. Från vänster Daniel Fröding, Sara Wird, Annelie Melander, Katarina Boman och Anders Ruuth.

FOTO: ROUF EKUNDO

Konsten att läsa en artikel om livskvalitetsforskning

Nu finns checklisten för att underlätta en kritisk värdering

|| Vilket värde kan livskvalitetsmått ha? Har observerade effekter klinisk mening?

Denna artikel syftar till att vägleda läsare av behandlingsforskning med särskild hänsyn till värdet av patientrapporterad information. Innehållet i artikeln bygger i stor utsträckning på erfarenheter från deltagande i internationella samlingsgrupper av läkare och metodexperter [1-4] samt i SBU-arbetsgruppen avseende systematisk kunskapssammanställning om behandling av astma och kroniskt obstruktiv lungsjukdom (KOL) [5, 6]. En grundläggande genomgång av livskvalitet som effektmått, med valda exempel från rapporten, återfinns i Läkartidningen 41/01 [7] och tjänar som inledning till det följande.

Aktuella riktlinjer vad gäller livskvalitetsmått sammanfattas här genom en checklista för kvalitetsbedömning av livskvalitetsforskning i kliniska studier [8].

|| Fråga 1

När kan livskvalitet som effektmått vara motiverat?

Svar: Vid sjukdomar där kliniska studier prövar alternativa behandlingsformer.

Kommentar: Livskvalitetsmått kan vara användbara för att värdera terapieffekter vid alla svåra, långvariga tillstånd. När bästa palliation är det främsta målet – dvs att göra patientens liv så bekvämt, funktionellt och tillfredsställande som möjligt – är också patienten själv den logiska experten på framgångsrik behandling [9].

Många goda exempel kan ges: när objektiva markörer av sjukdomsaktivitet inte finns (t ex vid migrän), när sjukdom karakteriseras av flera markörer (t ex astma), när sjukdom är lika med ett symtomkomplex (t ex irritabel kolon), när behandling förväntas förlänga liv men till priset av stora inskränkningar i livskvaliteten eller när behandling inte nämnvärt förväntas förlänga liv men kan öka livskvaliteten (t ex cancersjukdomar), när högeffektiv behandling ges svårt drabbade (t ex avancerad reumatoid artrit), när lågeffektiv behandling ges mindre allvarligt sjuka (t ex benign prostataförstoring), när sjukdom diagnostiseras via laboratorieprov utan åtföljande symtombörda (t ex högt blodtryck) eller när sjukdom klassificeras via ett enkelt mått men där sjukligheten är multipel (t ex svår fetma).

SAMMANFATTAT

Hälsorelaterad livskvalitet har under det senaste decenniet använts som effektmått vid kliniska prövningar i sådan utsträckning att det nu börjar lämpa sig som effektmått vid systematiska kunskapsöversikter.

En SBU-rapport om behandling vid astma/KOL har nyligen visat hur man genomför en systematisk kunskapsöversikt med hälsorelaterad livskvalitet som ett av de primära effektmåtten.

Särskilda kriterier behövs för kritisk granskning av patientrapporterad information med hänsyn till begrepps- och mätmodell, instrumentens tillförlitlighet och datas tolkningsbarhet.

Aktuella erfarenheter presenteras i en checklista för att underlätta läsning och värdering av artiklar om livskvalitetsforskning.

Evidensbaserad medicin

Vid jämförande kliniska prövningar med antagen lika verkningsgrad på längre sikt av två farmaka är det viktigt att dokumentera positiv inverkan på livskvaliteten (t ex åldrandets degenerativa sjukdomar).

Även vid akuta insjuknanden (t ex hjärtinfarkt, slaganfall) kan prövningar av optimal behandlingsarsenal inkludera livskvalitetsmått, utom vid behandling med otvetydig och väl-tolererad effekt [8].

Att kritiskt värdera en artikel om livskvalitetsforskning försvåras av det faktum att termen »livskvalitet« ökat mer i användning än i precision, dvs begreppet är inte alltid vetenskapligt definierat och mätt, och resultaten är därför svåra att uttolka. Även i ansedda medicinska tidskrifter finns en betydande variation. Detta blir tydligt i litteratursökningar, där

II Fakta 1

Psykometriska begrepp och definitioner för standardformulär avseende livskvalitet

Baserat på MOT-kriterierna i bearbetning [9, 10]:

Reliabilitet

Intern överensstämmelse (Cronbachs α);
Reproducerbarhet (test–retestkorrelation).

Validitet

Innehållsrelaterad (täckande, acceptabelt för patienterna, relevant för kliniker/andra vårdgivare);

Begreppsrelaterad (konvergent och diskriminant: »multitrait, multitrait–multimethod analysis«);

Kriterierelaterad (samtidig eller prediktiv: »known groups, known events analysis«).

Responsivitet

Känslighet för kliniskt viktiga förändringar (effektstorlekar: »effect size, standardised response mean«);

Tröskelvärden (betydelsefull skillnad: »minimal important difference«).

Tolkningsbarhet

Norm- och referensvärden (patient- och generella populationer/databaser);

»Number needed to treat«.

Genomförbarhet

Belastning för patienterna (företrädesvis självskattning);

Administrativa krav (intervju, kortversion).

Tillämpbarhet internationellt

Översättning/anpassning (begrepps- och språklig ekvivalens);

Psykometrisk kvalitet (jämförbar med originalinstrument).

t ex Medline ger ett kraftigt stigande antal träffar inom området »quality of life« från slutet av 1970-talet, då söktermen introducerades i Index Medicus. Med tillägget »questionnaires« till sökningen sjunker antalet artiklar drastiskt, och man är på god väg att ringa in den vetenskapligt grundade livskvalitetsforskningen i kliniska studier.

II Fråga 2

Vilka krav kan ställas på livskvalitet som effektmått?

Svar: Kriterier finns för bedömning av begrepps- och mätmodell, så att psykometrisk kvalitet kan säkerställas.

Kommentar: Ett pionjärbete har utförts av Medical Outcomes Trust (MOT) och dess Scientific Advisory Committee, en internationellt sammansatt expertgrupp. MOT är en icke-vinstdrivande organisation som garanterar tillgängligheten av goda resultatmått vad gäller hälsostatus/hälsorelaterad livskvalitet samtidigt som standardiseringen av innehåll, poängsättning, skalkonstruktioner och benämning skyddas. Syftet har varit att åstadkomma universell spridning av mätinstru-

ment som uppfyller särskilda krav på psykometrisk kvalitet [10]. Det bör noteras att konstruktörer av dylika mått inbjuds att ansöka om MOT-auktorisering; likvärdiga eller bättre instrument kan således föreligga utanför MOT-listan.

Verksamheten har dock bidragit till att skärpa uppmärksamheten så att standardiserade mått ökat i användning, även om ytterligare riktlinjer behövs för tolkning av resultaten från kliniska studier/läkemedelsprövningar. Här pågår ett utvecklingsarbete för mer enhetlig protokollskrivning och dokumentation. Incitamentet har varit forskarnas/klinikernas och läkemedelsindustrins behov att kommunicera kunskapsnivån till de myndigheter som fattar de avgörande besluten för framtidens farmakologiska behandlingsarsenal.

För avgränsning, operationalisering och exemplifiering av det mångtydiga begreppet hälsorelaterad livskvalitet (HRQL) hänvisas till Läkartidningen 41/01 [7]. Kärnpunkten är att standardformulär, som patienten själv kan hantera med kortfattad vägledning, omfattar väldefinierade nyckelområden kring symtombörda, funktion och välbefinnande – fysiskt, psykiskt och socialt. Livskvalitetsmått, generella såväl som specifika, är mångdimensionella, kvantitativa och uppbyggda enligt psykometrisk teori till sammansatta skalar, profiler och index med känd reliabilitet, validitet och känslighet för kliniskt relevanta förändringar (Fakta 1). De vanligast förekommande internationella formulären har anpassats och prövats för svenska förhållanden.

Kommentar till begreppen i Faktaruta 1

Att kritiskt värdera livskvalitetsdata ställer krav på läsarens kunskaper om och förståelse av psykometriska egenskaper hos frågeformulär. Följande uttolkning av begrepp och gränsvärden sammanfattar mer ingående beskrivningar [8-11] och skall ses som en första hjälp att skilja ut standardformulär från övriga. Så t ex duger inte angivelser av typen »livskvalitetsmåttet har god reliabilitet och validitet« utan understöd av empiriska data och referenser i metodbeskrivningen.

Reliabilitet: I en bra metodbeskrivning skall minst ett av följande uttryck för mätnoggrannhet vara angivet:

- Intern överensstämmelse mellan frågor i en skalkonstruktion bör ha överstigit 0,70 (korrelationsmetod vanligen Cronbachs alfa); ju fler frågor som ingår, desto högre alfa.
- Reproducerbarhet vid upprepad mätning (test–retest) bör ha överstigit 0,70 (parametrisk metod) vid stabil sjukdom och en till fyra veckors intervall mellan mätningarna. Om distributionsfria metoder används, t ex kappas, bör gränsen 0,60 ha passerats. För frågeformulär som administrerats via intervju bör reproducerbarheten vara angiven såväl för enskilda intervjuare som mellan intervjuare.

Validitet: Följande tre begrepp bör finnas i en bra metodbeskrivning. Det bör noteras att angivelser om validitet varierar betydligt i vetenskapliga publikationer; en minsta acceptabel nivå är därför svårare att fastställa än när det gäller mätnoggrannheten.

- Innehållsvaliditet, »face validity«, dvs att frågeformuläret täcker hela området på ett övertygande sätt, klart och enkelt utan ett överflöd av frågor bör vara beskrivet i termer av hur relevant, trovärdigt och acceptabelt det framkommit i pilotstudier av patienter, ibland även av läkare, andra vårdgivare och anhöriga, s k fokusgrupper.
- Begreppsvaliditet, dvs att formulärets inre struktur är känd och att mätningarna korrelerar med liknande mått på ett förväntat sätt, skall vara testad och redovisad. Jämförelser

II Fakta 2

Några vanliga begrepp inom psykometri

Mätning

Skattning, kvantifiering. Att med viss metodik (instrument) tillordna poäng till kvantiteter av en egenskap.

Instrument

Frågeformulär, enkät. Oftast avses standardformulär, dvs formulär som testats enligt psykometrisk teori och bestämda kriterier (Fakta 1).

Domän

Avser ett särskilt område av egenskaper. De frågor, »items«, som ingår skall var och en omfatta olika aspekter av egenskapen i fråga. Ett instrument kan bestå av en eller flera domäner.

Skala

Mätning av domän. Antalet frågor i en skala varierar. Kan vara en enstaka fråga, t ex VAS (visuell analog skala), där svarsalternativen är lika med skalan, »single-item scale«. Oftast menas enligt psykometrisk teori skalpoäng, dvs en summering av svarsalternativ från ett antal frågor, »multi-item scale«.

Profil

Grafisk beskrivning av instrumentets skalpoäng.

Index

Summakomponent, skala av 2:a ordningen, en summering av ett antal skalpoäng.

II Fakta 3

Vanliga metoder för beräkning av responsivitet hos livskvalitetsmått

Effektstorleksberäkning (ES) innebär att skillnaden i medeltal mellan måttillfällen (t ex mellan baslinje och uppföljningstillfälle) divideras med ett spridningsmått (t ex standardavvikelsen vid baslinjen, »effect size«).

Denna formel kan direkt överföras till jämförelse mellan grupper (dvs skillnad i medeltal mellan grupper dividerad med den hopslagna standardavvikelsen för grupperna).

»**Standardized response mean (SRM)**» innebär att skillnaden mellan två måttillfällen divideras med standardavvikelsen för denna skillnad i en grupp. Formeln kan direkt generaliseras till jämförelse mellan grupper (dvs skillnaden i medeltal mellan två måttillfällen i grupperna dividerad med den hopslagna standardavvikelsen för denna skillnad).

Tumregler för tolkning av ES eller SRM [12]:

Trivial <0,20

Liten 0,20–0,50

Moderat 0,50–0,80

Stor >0,80

kan ha gjorts med tidigare kända frågeformulär för att stärka hållbarheten. En mätning skall ha visat starkt samband med andra mått på samma begrepp. Lika viktigt är att mätningen visat mindre starkt samband med mått på andra, kanske närliggande, begrepp. Ett instruments olika frågor och domäner (Fakta 2) skall ha genomgått åtminstone en grundläggande analys av den inre strukturen, »multitrait analysis«. Varje fråga skall ha visat starkt samband med den egna domänen (>0,40 justerat värde), svagare med alla andra domäner (< två standardavvikelser för korrelationsmatrisen). Om mer än en metod har använts vid datainsamlingen eller skalkonstruktionen bör en utvidgad analysmetod ha använts i tillägg, »multitrait-multimethod analysis«. En sådan korrelationsmatris innehåller interkorrelationer mellan två eller flera begrepp som mäts med två eller flera metoder. En mätning skall ha visat signifikant starkare samband med andra mätningar till vilka den bör vara relaterad (konvergent validitet; korrelation mellan 0,30 och 0,70) än med andra mätningar som den inte bör vara relaterad till (diskriminant validitet; korrelation <0,30).

- Kriterievaliditet, dvs att ett frågeformulär har ett värde över en bestämd nivå vid jämförelse med en vald »gyllene standard« av det begrepp som mäts, skall vara dokumenterad. Skalpoäng för olika delar av frågeformuläret skall ha visats relaterade till ett eller flera externa mått på klinisk validitet, t ex självskattad gångförmåga versus mätning av 6 minuters gångsträcka på klinik, eller självskattad depression versus klinisk depression. Analysen kan gälla samtidig eller prediktiv validitet (korrelationsvidd 0,30–0,70) eller förmågan att skilja mellan distinkta grupper i klinisk mening, t ex patienter med olika svårighetsgrad av sjukdom, »known groups analysis«.

Responsivitet/känslighet för förändringar: Denna egenskap är livsviktig för ett frågeformulärs användbarhet i kliniska studier. Instrumentets förmåga att upptäcka förändringar som bedöms kliniskt viktiga, även om förändringarna är små, skall vara prövad och dokumenterad på mer än ett sätt. Redovisning av en signifikant förändring i skalpoäng efter en intervention räcker inte. Livskvalitetsdata över tid bör förete samvariation med andra mått på förändring enligt i förväg uppställda hypoteser.

I det följande beskrivs ytterligare två sätt att uppskatta betydelsen av förändring: distributions- och ankarbaserad metodik (Fakta 3, Tabell I).

- Bedömning av effektstorlek (distributionsbaserad beräkning) gäller vikten av en funnen skillnad för skalor i samma instrument över flera studier, eller för skalor från olika instrument med varierande poängsättning och variationsvidd, dvs hur bra man kunnat särskilja »signalen« från »bruset«. Metodiken förekommer alltmer i vetenskaplig redovisning av livskvalitetsmått för en mer enhetlig bedömning (Fakta 3). Så t ex kunde vi dra slutsatser med högt bevisvärde i en systematisk genomgång av livskvalitet efter fetmabehandling (kommande SBU-rapport), där stor effekt (>0,80) kunde knytas till flera mått på fetmarelaterade problem, funktion, allmän hälsa och välbefinnande förutsatt en stor och bibehållen vikttnedgång på lång sikt. I befolkningsstudier från olika länder med samma instrument, SF-36, kunde vi genom att beräkna effektstorlekar för delskalorna bättre tolka vilka aspekter av livskvaliteten som sjunker mest med ökande fetma.
- Bedömning av minsta betydelsefulla skillnad (ankarbaseerad beräkning) stöder sig på information om tröskelvärden

Tabell I. Skattning av minsta betydelsefulla skillnad i livskvalitet med globalskattad förändring som standard.

Svar på global skattning av förändring ¹	Sämre	Bättre	Tolkning av förändring	Förändring i medeltal i livskvalitetsskala
En mycket stor del	-7	+7	Stor	1,5
En stor del	-6	+6		
En hel del	-5	+5	Måttlig	1
Måttligt	-4	+4		
Lite	-3	+3	Liten	0,5
Mycket lite	-2	+2		
Nästan detsamma	-1	+1	Ingen	
Ungefär detsamma	0			

¹ »På det hela taget, har du upplevt någon förändring i andfåddhet under de dagliga aktiviteterna sedan du besökte oss senast?«

II Fakta 4

Kulturell och språklig anpassning av formuläret

»Poolad« första översättning. Minst två oberoende översättningar från originalspråk till det nya språket, som förenas till en så god första översättning som möjligt.

Återöversättning. Oberoende översättning av versionen ovan till originalspråket av en eller två personer.

Testversion efter jämförelser av versionerna. En diskussion mellan översättarna och expert- och lekmannapaneler. Särskilt viktigt att få god läsbarhet, enkelt språk, begreppsmässig likhet men med språkbruk och exempel som människor lätt kan relatera till. Revisioner ofta nödvändiga.

Pilottestning. Testning med patienter (fokusgrupper) som instrumentet är avsett för, i syfte att fånga upp nya kulturella aspekter som inte varit aktuella i ursprungslandet. Fokus på innehållsvaliditet.

Fältstudie. Utprovning för att testa om instrumentet motsvarar originalet. Uppläggnings för testning av reliabilitet och olika aspekter av validitet.

Slutversion. Alla eventuella skillnader mellan original och översatt/anpassat formulär skall förklaras och värderas. Håller det nya formuläret samma kvalitet som originalet?

från studier där minsta kliniskt meningsfulla förändring värderats. Sådan information efterfrågas allt oftare i kliniska prövningar och börjar synas i litteraturen, även om diskussionen kring bästa tillvägagångssätt fortsätter. Detta återknyter till svårigheterna att fastställa ett frågeformulärs kriterievaliditet. I Tabell I beskrivs en enkel metod där patienterna fått skatta upplevd förändring i andfåddhet över en viss tid (enstaka fråga, 1,5 skalsteg för gradering av försämring respektive förbättring från -7 till +7). Samma patienter har vid upprepade tillfällen besvarat ett standardformulär om livskvalitet (flera frågor om olika aspekter av livskvaliteten, sammanförda till en summapoäng, skalvidd 1-7). En förändring i medeltal på 0,5 poäng till det bättre eller sämre kan enligt detta tillvägagångssätt tolkas som en minsta betydelsefull skillnad [13]. Metoden är beroende av flera studiespecifika karakteristika, och resultatet kan inte utan vidare generaliseras.

Tolkningsbarhet: De distributions- och ankarbaserade uttrycken för livskvalitetseffekter enligt ovan ger komplette-

rande tolkningsunderlag men finns inte alltid beskrivna. Ytterligare sätt att definiera betydelsen av behandlingseffekter innefattar hur många personer som behöver behandlas under en viss period för att en person skall förbättra sin livskvalitet upp till gränsen för betydelsefull skillnad. Det antal som når gränsen i behandlingsgruppen korrigerat för motsvarande antal i kontrollgruppen blir »number needed to treat«, NNT [14]. Detta efterfrågas numera vid värdering av kliniska prövningar men har hittills sällan beskrivits i vetenskapliga publikationer.

En god tolkningshjälp erhålls genom att relatera livskvalitetsdata från kliniska studier till motsvarande nivåer i olika väldefinierade grupper (patient- och normalpopulationer). En angivelse av ett eller flera referensvärden bör finnas i varje publikation av livskvalitetsdata. Bestämning av normal nivå och tröskelvärden för markerad inskränkning av olika aspekter av livskvaliteten kan därmed göras och underlätta tolkningen av behandlingseffekter. Jämförelser med andra kliniska effektmått kan också bidra till värdering av livskvalitetsdata.

Av särskild vikt för värdering av livskvalitet som effektmått är att hanteringen av databortfall beskrivits. Det bör gälla både partiellt bortfall i besvarade formulär och totala antalet missade formulär över tiden. De metoder som använts för att ersätta databortfall skall vara motiverade bl a av huruvida bortfallen bedömts som slumpmässiga eller systematiska.

Genomförbarhet: Praktikaliteter skall finnas i metodbeskrivningen för att läsaren själv skall kunna värdera frånvaro av bias i resultaten. Här sammanfattas de viktigaste faktorerna i livskvalitetsstudier, dvs patientbörda och administrativ belastning:

- Kraven på patienten skall vara beskrivna, då risken för bias minskas om formuläret är lätt att fylla i och rimligt kortfattat; för kliniska prövningar kan genomsnittstiden 1,5 minuter vara ett riktvärde. En äldre eller svårt sjuk person behöver ofta assistans, vilket skall ha angetts. I speciella studier (t ex vid slaganfall) kan anhörig- eller vårdgivarskattningar accepteras som närmevärden, förutsatt att de använts under hela undersökningsperioden.
- Administreringen av formulär är viktig; instruktionerna bör innehålla en beskrivning av de särskilda resurser som avsatts – tidsmässigt, lokalmässigt, kunskapsmässigt – för att säkra datakvaliteten. Enhetlighet i när, var och hur respondenterna fått frågeformulären skall framgå, eftersom detta har stor betydelse när det gäller att undvika onödiga felkällor och bortfall av data. Ett eventuellt byte av inter-

II Fakta 5

Checklista för kvalitetsbedömning

(baserat på ERIQA Group [8])

Är värdet av HRQL-mätningen och valet av instrument adekvat motiverat?

Är relevansen att mäta HRQL motiverad vid sjukdomen i fråga?

Är syftet med HRQL-mätningen klart formulerat?

Är valet av HRQL-mått dokumenterat?

Täcker formuläret de domäner där förändring är förväntad?

Är specificerade, testbara HRQL-hypoteser uppställda?

Finns det en klar beskrivning i designen av de delmoment som angår HRQL-mätningen?

Är HRQL definierat som primärt eller sekundärt effektmått?

Är urvalet representativt för studiepopulationen och populationen patienter som sannolikt kommer att få behandlingen?

Urval av centra (vid multicenterstudier).

Inklusions- och exklusionskriterier.

När och hur ofta utförs HRQL-mätningarna?

Hur och var administreras HRQL-mätningarna?

Belastning för patienten: tidsåtgång.

Datamonitorering och kvalitetssäkring.

Procedurer för minimering och hantering av bortfall av data.

Är de psykometriska egenskaperna hos HRQL-måtten adekvat beskrivna och dokumenterade?

Antal frågor och domäner; typ av skala och poängberäkning.

Reliabilitet (intern överensstämmelse, reproducerbarhet).

Validitet: innehålls-, begrepps-, kriterievaliditet.

Responsivitet/känslighet för förändringar.

Språklig/kulturell anpassning.

Styrkt psykometrisk kvalitet för översatta formulär.

Finns det en detaljerad statistisk analysplan av HRQL-data beskriven?

Överlägsenhet eller jämförbarhet som hypotes.

Urvalsstorlek och statistisk styrka; »intent to treat«-analys.

Beskrivande och hypotesprövande statistik för värdering av förändringar över tid och skillnader mellan grupper.

Procedurer för att hålla det totala typ I-felet på en vald nivå (t ex 0,05) i fallet med många testningar.

Ersättning av bortfall av data.

Är rapporteringen av HRQL-resultat detaljerad och informativ?

Information om studiepopulationen: Deltagandefrekvens (vid baslinjen och uppföljningstillfällena); demografiska och medicinska karakteristikerna av HRQL-studiens population; datas fullständighet (dvs värderingsbara formulär och skalor).

Presentation av resultaten: Fullständig redovisning av skalor och index; fördelning av HRQL-värden inom och mellan grupper; 95 procents konfidensintervall av skillnad och/eller oddskvot.

Kan HRQL-resultaten tolkas i termer av klinisk relevans?

Effektstorlek och/eller »standardised response mean«; jämförelser av HRQL-värden med normvärden/andra liknande studier/populationer; samband med kliniska och andra externa kriterier i överensstämmelse med förväntad nivå.

Tröskelvärden för minsta betydelsefulla skillnad.

»Number needed to treat«.

vjuformulär till självskattningsformat, och vice versa, skall ha föregåtts av särskild prövning av instrumentets egenskaper i alla ovan nämnda psykometriska kriterier.

Tillämpbarhet internationellt: Kulturella och språkliga anpassningar av frågeformulär från originalspråket till ett annat språk följer en bestämd sekvens av åtgärder med variation i grad av sofistikerade insatser. Redovisning skall finnas om både metoder som använts och utfallet. I Fakta 4 beskrivs en vanligen använd sekvens av åtgärder för godtagbar ekvivalens mellan det anpassade formuläret och originalet.

II Fråga 3

Är ett sjukdomsspecifikt livskvalitetsmått tillräckligt för säkerställd effekt?

Svar: Adekvat begrepps- och mätmodell avgörande (jämför generella mått), dvs mångdimensionellt mått (HRQL) och psykometrisk kvalitet.

Kommentar: Det viktigaste kriteriet för ett specifikt livskvalitetsmått är frågan hur säkert det förmår täcka väsentliga aspekter bundna till ett givet hälsoproblem. Goda exempel återfinns i SBU-granskningen av behandling vid astma/KOL [5, 6]. Det kan här vara fruktbart att nämna exempel på mätningar som inte uppfyller kriterierna för HRQL-mått. Mätning av en enskilda domän, en diagnostisk skala, gradering av svårighetsgrad av sjukdom/symtom, checklistor avseende närvaro/frånvaro av symtom och listning av bieffekter av behandling utgör sådana exempel. Funktionsmått eller symtomskalor som mäter t ex trötthet, smärta, ångest eller depression ingår ofta som viktiga delar i HRQL-mått. Inslag av sådana enskilda mått gör dock inte namnet livskvalitetsstudie berättigat.

II Avslutande kommentar

Denna artikel avser att vägleda läsaren av livskvalitetsforskning i kliniska studier. En checklista (Fakta 5) sammanfattar viktiga punkter för att underlätta en kritisk värdering. Man bör här komma ihåg livskvalitetsområdets relativt korta historia inom klinisk medicin, där studieprotokoll först på senare tid professionellt planerat för patientrapporterade mått. Publikationsreglerna har inte heller varit strikta, utan stor variation föreligger avseende presentation av t ex frågeformulärens psykometriska egenskaper, datakvalitet, analys av resultat och tolkning av fyndens betydelse. Med ökande vetenskaplig standard och fastare riktlinjer för tolkning spås livskvalitetsmått få stort värde i medicinskt beslutsfattande [15].

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Aaronson N, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez N, et al. For the EORTC Study Group on Quality of Life. The EORTC QLQ-C30: a quality of life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365-76.
2. Bullinger M, Alonso J, Apolone G, Leplège A, Sullivan M, Wood-Dauphinee S, et al. For the IQOLA Project Group. Translating health status questionnaires and evaluating their quality. The IQOLA-Project approach. *J Clin Epidemiol* 1998;51:913-23.
3. Acquadro C. For the ERIQA Group. The European Regulatory Issues on Health-Related Quality of Life Assessment (ERIQA) Project: Preliminary results and plans for the future. *Quality of Life Newsletter* 2000;24:1-3.
4. Staniek V. Cochrane Health-Related Quality of Life Methods Group (HRQoL MG): toward registration. *Quality of Life Newsletter* 2001;26:1-2.
5. Behandling av astma och KOL. En systematisk kunskapsman-

- ställning. Stockholm: SBU, Statens beredning för medicinsk utvärdering; 2000. Rapport nr 151.
6. SBU-rapport om astma och kroniskt obstruktiv lungsjukdom. Systematisk kunskapsöversikt som underlättar behandlingsvalet. *Läkartidningen* 2001;98:288-93.
 7. Sullivan M, Tunsäter A. Hälsorelaterad livskvalitet informativt effektmått i kliniska studier. Tillämpningsexempel: astma/kroniskt obstruktiv lungsjukdom. *Läkartidningen* 2001;98:4428-33.
 8. Chassany O, Sagnier P, Marquis P, Fullerton S, Aaronson N. For the European Regulatory Issues on Health-Related Quality of Life Assessment (ERIQA) Group. Patient reported outcomes and regulatory issues: the example of health-related quality of life. *DIA Journal* 2002;36:209-38.
 9. Sullivan M, Karlsson J, Taft C. How to assess quality of life in medicine: Rationale and methods. In: Guy-Grand B, Ailhaud A, editors. *Progress in obesity research: 8*. London: John Libbey; 1999. p. 749-55.
 10. The Medical Outcomes Trust: Scientific Advisory Committee. Instrument Review Criteria. *MOT Bulletin* 1995;3(4):1-4.
 11. Apolone G, De Carli G, Brunetti M, Garattini S. Health-related quality of life (HR-QOL) and regulatory issues. An assessment of the European Agency for the Evaluation of Medicinal Products (EMA). Recommendations on the use of HR-QOL measures in drug approval. *Pharmacoeconomics* 2001;19:187-95.
 12. Cohen J. *Statistical power analysis for the behavioral sciences*. New York: Academic Press; 1978.
 13. Jaeschke R, Singer J, Guyatt GH. Measurement of health status: ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10:407-15.
 14. Guyatt GH, Juniper EF, Walter SD, Griffith LE, Goldstein RS. Interpreting treatment effects in randomised trials. *BMJ* 1998;316:690-3.
 15. Higginson IJ, Carr AJ. Using quality of life measures in the clinical setting. *BMJ* 2001;322:1297-300.

SUMMARY

The art of reading an article on quality of life research – new check list available to facilitate a critical evaluation

Marianne Sullivan

Läkartidningen 2002;99:2933-8

During the past decade health-related quality of life has been applied in clinical trials in many chronic diseases to such an extent that it is now feasible as outcome in systematic reviews. A recent report on treatment of asthma and chronic obstructive pulmonary disease showed how an evidence based review could be completed with health-related quality of life as one of the primary outcome measures. Thus, patient-reported outcome of treatment efficacy is nowadays adequately standardised for scientific reviews according to the conventional criteria of evidence based medicine. In addition, specific criteria are needed with reference to conceptual and measurement models, psychometric standard of instruments and interpretability of results. Current experience is presented in a check list to facilitate reading and evaluation of articles in quality of life research.

Correspondence: Marianne Sullivan, Health Care Research Unit, Dept of Internal Medicine, Sahlgrenska Universitets-sjukhuset, SE-413 45 Göteborg, Sweden
(marianne.sullivan@medicine.gu.se)

Särtryck

Läkartidningen

Vad händer i slutna rum då läkare blir problem? Hur löser man konflikter vid missöden i vården? Vem kan ge ett bra stöd till anmälda läkare?

Under 1994–95 publicerade *Läkartidningen* 21 artiklar om problemläkare, läkarproblem, ansvarsfrågor och patientförsäkringen.

De har nu samlats i ett 80-sidigt häfte, som kan beställas med kupongen nedan.

Priset är 65 kronor



Missöden misstag missbruk

Beställer härmed.....ex
av "Missöden, misstag, missbruk"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till **Läkartidningen**
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Per Carlsson, professor i utvärdering av medicinsk teknologi, Centrum för utvärdering av medicinsk teknologi, Linköpings universitet, föreståndare för Nationellt kunskapscentrum för prioriteringar inom vård och omsorg, Linköping (*perca@cmt.liu.se*)

Evidensbaserad vård förutsätter kunskap om kostnadseffektivitet

|| Hälsoekonomi har under de senaste tio åren utvecklats från att tidigare ha varit ett udda ämne, som en liten grupp av företrädesvis ekonomer intresserade sig för, till en relativt väl-etablerad disciplin. Den senaste världskonferensen i hälsoekonomi, i engelska York, samlade över 1 200 delegater. De fick bl a lyssna till Nobelpristagaren Amartya Sen. Hans föredrag, liksom flera andra bidrag, kretsade kring den centrala frågan om vad som menas med jämlik hälsa. Är det tillräckligt att tillgången på sjukvårdsresurser är jämlik eller krävs något mer? Svaret är förstås nej och ja på dessa frågor. Ändå handlar den sjukvårdspolitiska debatten nästan bara om fördelning av och tillgång till hälso- och sjukvårdens resurser.

Begreppet hälsoekonomi står för både ett förhållningssätt och en tillämpning av ekonomiska teorier, metoder och synsätt på hälso- och sjukvården. Förhållningssättet handlar om att tillägna sig ett ekonomiskt tänkande som bl a innebär att resurserna i samhället är begränsade och att det inte går att tillfredsställa alla mänskliga behov med tillgängliga resurser, utan att vi som kollektiv eller enskilda individer ställs inför val. Kravet på hushållning av begränsade resurser gäller för sjukvården på samma sätt som inom andra samhällssektorer, liksom för den egna privatekonomin. Detta är på en övergripande nivå knappast kontroversiellt. Ingen som är verksam i vården anser väl att det är försvarbart med resursslöseri utan tycker nog att tillgängliga resurser i form av personal, material, lokaler m m skall användas så förnuftigt som möjligt.

På det individuella planet finns det dock en rad konflikter. I relationen till enskilda patienter förväntas läkaren eller sjuksköterskan vara patientens agent och enbart ha patientens bästa för ögonen, och underförstått inte ta hänsyn till andras behov. I praktiken sker inte detta fullt ut, och det är inte heller rimligt att så sker. Läkare och annan vårdpersonal tvingas både i det dagliga värdarbetet och när man agerar i ledande ställning göra avvägningar mellan olika patienters och patientgruppers behov av vård, förväntad nytta och kostnader. Fördelningen av tillgängliga resurser handlar inte bara om att tillfredsställa de behov som visar sig på mottagningen utan man måste också ställa frågan om det finns behov av insatser i befolkningen som inte efterfrågas. De finns grupper med låg autonomi som behöver extra stor uppmärksamhet.

Hälsoekonomiska utvärderingar ökar i betydelse

De flesta personer som är verksamma i vården kommer i första hand i kontakt med den mest tillämpade delen av hälso-

SAMMANFATTAT

När inte alla mänskliga behov kan tillfredsställas ställs vi inför valsituationer.

Kostnadseffektivitetsprincipen innebär att eftersträva en rimlig relation i vården mellan kostnader och patientnytta.

Kraven ökar på att redovisa kostnadseffektivitet inför beslut om läkemedel.

Modellstudier är inte substitut till välgjorda kliniska studier utan ett sätt att systematiskt sammanställa bästa tillgängliga information i avvaktan på säkrare data.

Hälsoekonomiska resultat skall lika lite som annan klinisk forskning användas oreflekterat vid beslut som rör enskilda patienter.

Evidensbaserad medicin

ekonomin, nämligen utvärderingar. Sådana görs med syftet att förbättra effektiviteten i vården även om hälsans fördelning också är en viktig aspekt att ta hänsyn till. Idag finner man hälsoekonomiska artiklar i de flesta medicinska tidskrifter som är inriktade på klinisk forskning. Men det är trots allt vid utvärdering av nya läkemedel som hälsoekonomi blivit mest etablerad. I flera länder ställs numera krav på producenterna av nya läkemedel att redovisa kostnadseffektivitet inför beslut om subventionering, vilket bidragit till en ökad tillgång på hälsoekonomiska studier. I samband med att den svenska läkemedelsförmånsnämnden kommer att börja sitt arbete den 1 oktober 2002 så kommer hälsoekonomiska värderingar att spela en allt viktigare roll även i vårt land. Med andra ord är det inte längre en fråga om huruvida hälsoekonomiska aspekter skall beaktas vid beslut som rör användning av resurser i sjukvården utan hur kunskapen skall användas på bästa sätt.

Även om tillgången på hälsoekonomiska utvärderingar

ökat under senare tid och metodiken blivit mer standardiserad kvarstår det en mängd frågor om hur hälsoekonomiska data skall hanteras i beslutsfattandet. Det gäller bland annat hur kvalitetsgranskning och värdering av resultaten skall ske. Beslut om införande av nya läkemedel eller nya tekniker fattas ofta under stor osäkerhet och då det vetenskapliga underlaget är mycket begränsat. När man vill göra en bedömning av den förväntade patientnyttan och ställa den i relation till kostnaderna i samhället måste man i regel basera sin analys på antaganden.

En annan aktuell fråga gäller hur man skall använda kunskapen om kostnadseffektivitet vid prioriteringar i vården. Prioriteringsutredningen och den efterföljande riksdagsbehandlingen skapade stor förvirring när man skrev att kostnadseffektivitetsprincipen är underordnad behovsprincipen utan att utförligt diskutera vad det kan innebära. Regeringen har emellertid i proposition 2001/02:63 bättre än tidigare klargjort vilken roll kostnadseffektivitet skall spela vid bedömning av nya läkemedel. Enligt regeringen innebär kostnadseffektivitetsprincipen att en rimlig relation mellan kostnader och effekt, mätt i förbättrad hälsa och livskvalitet, bör eftersträvas vid val mellan olika verksamheter och åtgärder.

Vad menas med kostnadseffektivitet?

Kostnad är en kvantifiering av en resurs som åtgår för att åstadkomma till exempel en behandling, undersökning eller ett samtal med anhöriga. Här kan ingå delkostnader för personal, laboratorieanalyser och läkemedel. Grova kostnadsbedömningar görs mer eller mindre intuitivt vid alla former av medicinskt beslutsfattande eller omvårdnadsbeslut. Det uppfattas till exempel inte som ändamålsenligt att med magnetkamera undersöka alla som söker för huvudvärk men som teoretiskt skulle kunna ha ett allvarligt tillstånd. Ibland är besluten och deras konsekvenser enkla att överblicka och kräver inga sofistikerade hälsoekonomiska analyser. Det kan handla om läkarens egen tidsinsats i en situation där den förväntade patientnyttan är uppenbar. Andra gånger är beslutsproblemen mer komplicerade.

Inför ett ställningstagande om ett nytt program för till exempel screening för tidig upptäckt av cancer går det knappast att enbart på basis av egna erfarenheter bedöma om den förväntade hälsovinsten står i rimlig proportion till de framtida kostnaderna för programmet. Det krävs att nyttan och riskerna med screeningen studerats i kontrollerade studier liksom att de kostnader som uppstår i olika delar av samhället måste analyseras på ett vetenskapligt godtagbart sätt. Så snart man tvingas att välja hur man vill använda sina resurser för att uppnå så mycket som möjligt av det man önskar uppnå har man behov av information om konsekvenserna av olika handlingsalternativ. Det är här som hälsoekonomiska metoder kommer in bilden. Sådana kan användas för att identifiera, kvantifiera och värdera kostnader för och nytta av olika handlingsalternativ.

Vid hälsoekonomiska utvärderingar jämförs kostnadseffektiviteten hos minst två alternativa strategier. Det kan gälla två operationsmetoder, två läkemedelsbehandlingar eller en jämförelse mellan en medicinsk behandling och ett alternativ med omvårdnadsinsatser. Kostnadseffektivitetsmätningen är i regel en kvot mellan en primär effektvariabel och kostnad, som kronor per vunnet levnadsår eller kronor per vunnen smärtfri dag.

Under årens lopp har jag mött läkare som ifrågasatt hälso-



ekonomiska beräkningar därför att de tyckt att det är olämpligt att använda ekonomiska kalkyler vid beslut som rör enskilda patienter. Jag och andra hälsoekonomer har ingen annan åsikt. Det har aldrig varit tänkt att informationen ska användas oreflekterat vid beslut som rör enskilda patienter lika lite som annan patientnära klinisk forskning är giltig som enda beslutsunderlag för behandlingsbeslut för enskilda patienter. Ändå behövs det ibland kliniska studier som visar den förväntade hälsoeffekten eller kostnaden av en åtgärd som referens att utgå från. Det är från sådana underlag behandlingsrekommendationer och riktlinjer kan tas fram som beslutsstöd. I det enskilda behandlingsbeslutet måste givetvis den enskilde patientens eller patientgruppens specifika behov och egenskaper och önskemål vägas in.

Krångligt att beräkna kostnaden

Vid hälsoekonomiska utvärderingar är kostnadsbegreppet centralt. För övrigt beräknas kostnaderna på samma sätt oavsett vilken av de nedan beskrivna metoderna som används. När läkaren fattar ett behandlingsbeslut tas alltid resurser i anspråk som i princip skulle kunna användas för något annat ändamål. Alternativkostnaden, det vill säga värdet av den hälsovinst som skulle uppkommit vid bästa tänkbara alternativa resursanvändning, är mest korrekt att använda i samhällsekonomiska kalkyler. I praktiken är det inte möjligt att ta reda på detta, utan kostnadsberäkningarna vid hälsoekonomiska utvärderingar måste göras på ett annat sätt.

Kostnader för medicinska åtgärder vid diagnostik och behandling försöker man istället beräkna med hjälp av marknadspriser och utifrån sjukvårdens egna kostnadsredovisningar. Beräkning av kostnader för specifika åtgärder i sjukvården är fortfarande både komplicerad och tidsödande eftersom kostnadsredovisningen inte är utformad på ett optimalt sätt. Vid vetenskapliga studier är det vanligt att kostnadsuppgifter för olika åtgärder, till exempel laboratorieanalyser, röntgenundersökningar och mottagningsbesök, inhämtas från flera sjukhus för att undvika extrema värden [1].

Vilka kostnader skall tas med i hälsoekonomiska utvärderingar? Svaret är att alla kostnader som bedöms relevanta skall ingå. Vad som kan anses relevant beror i sin tur på den frågeställning som studien skall besvara. Om man enbart är intresserad av att svara på frågan hur en ny operationsmetod påverkar klinikens ekonomi så är det acceptabelt att begränsa sin datainsamling till de kostnadsslag som är direkt relaterade till värden vid den aktuella kliniken. Vid hälsoekonomiska utvärderingar rekommenderas dock ett samhällsperspektiv. Det innebär att man försöker kartlägga alla relevanta kostnader och effekter oavsett var de uppstår i samhället och oavsett när de inträffar i tiden. Det innebär att kostnader liksom besparingar och andra effekter som inträffar även långt fram i tiden skall tas med. Framtida kostnader och effekter bör emellertid ges ett lägre värde än nuvärdet (diskontering). I regel görs en nedskrivning på 3–5 procent per år. Diskontering har sin grund bland annat i att man vill beakta den osäkerhet som finns om framtiden (jämför med uttrycket: »Hellre en fågel i handen än tio i skogen«). För en utförligare förklaring hänvisas till läroböcker i ämnet [2].

Andra åtgärder än medicinska, som transporter till och från vårdinrättningen, beräknas som den faktiska kostnaden för egen bil eller kostnaden för taxiresor. Den här typen av kostnader liksom tidskostnader skall tas med om de bedöms ha be-

tydelse för resultatet. Vid utvärdering av till exempel en hälsokontroll till vilken i huvudsak friska personer inbjuds är det angeläget att såväl produktionsbortfall som andra tidskostnader kartläggs och värderas. När man vill beräkna kostnadseffektiviteten hos olika strategier för att bedriva hemsjukvård är det istället viktigt att kartlägga anhörigas tidsinsats för de behandlingsalternativ som jämförs. Metoder för värdering av produktionsbortfall liksom bestämning av värdet av anhörigas tid är en källa till diskussion. Här saknas av förståeliga skäl en handfast metodik eftersom förhållandena skiljer sig kraftigt mellan olika vårdssituationer. En viktig princip vid hälsoekonomiska studier är dock att alla resurser som identifierats i form av till exempel läkartid, operationsminuter, EKG-undersökningar och anhörigas tidsinsats redovisas i fysiska storheter som timmar, antal undersökningar, dagar m m. Härigenom kan läsaren av studien såväl följa hur man kommit fram till resultatet som göra egna beräkningar med hjälp av lokala kostnadsuppgifter. Även relevanta »kostnader« som man anser inte går att skatta i kronor bör anges om de inte fångas med hjälp av något livskvalitetsmått.



Vidare skall indirekta kostnader i form av patientens tidskostnader beaktas. Om det är vanligt att ta med framtida kostnader som direkt relaterade till sjukdomen är det mindre vanligt att ta med kostnader för orelaterade sjukdomar eller konsumtion för mat, kläder m m som uppstår vid vunna levnadsår. Ekman och medarbetare har gjort sådana beräkningar i utvärdering av läkemedlet bisoprolol vid behandling av kronisk hjärtsvikt som tidigare publicerats i Läkartidningen [3]. Det finns teoretiska fördelar med detta angreppssätt men också nackdelar. Om man har ambitionen att ta med kostnader för orelaterade sjukdomar och samtliga konsekvenser på samhällsekonomin kan resultaten bli osäkra och svåra att tolka då det är metodologiskt besvärligt att beräkna dessa kostnader på ett invändningsfritt sätt. Jämförelser mellan studier från olika länder blir också svårare att göra.

Olika typer av hälsoekonomiska studier

Det finns olika typer av hälsoekonomiska utvärderingsmetoder. De har vissa likheter, särskilt på kostnadssidan, men också vissa olikheter som är värda att känna till. Kostnadsminimeringsanalys (cost-minimization analysis) förekommer som ett specialfall och kan användas i de fall den medicinska effekten är lika mellan de studerade alternativen. I sådana fall brukar man nöja sig med att jämföra enbart kostnaderna. Det alternativ som har den lägsta totala kostnaden är att föredra.

Vanligtvis är effekten olika mellan alternativen. Ofta innebär ett nytt läkemedel eller en ny operationsmetod att en något bättre effekt uppnås men till högre kostnad. Kostnadseffektanalys (cost-effectiveness analysis) syftar till att relatera extrakostnaden (marginalkostnaden) för det dyrare alternativet med vad man vinner i extra hälsoeffekt (marginalnyttan). Vad detta kan innebära i praktiken framgår av följande exempel från SBU Alert [4].

Ett relativt nytt läkemedel, klopidogrel, som är godkänt för behandling av symtomgivande aterosklerotisk sjukdom har prövats i en stor undersökning istället för ASA. Enligt studien minskar den relativa risken för en allvarlig hjärt-kärlhändelse med 8,7 procent. I absoluta tal minskar risken med 0,5 procent. För att undvika en händelse måste alltså cirka 200 patienter behandlas under ett år till en merkostnad av 6 000 kronor per patient. Ett preliminärt kostnadseffektmått är att det kostar 1,2 miljoner kronor för att undvika en allvarlig händelse. Om denna kostnad står i rimlig relation till hälsovinsten är dels en värderingsfråga, dels en empirisk fråga. För att analysen skall bli fullständig måste kostnader för att behand-

la slaganfall, hjärtinfarkt m m tas med, liksom de besparingar som görs tack vare att sådana händelser kan undvikas. Dessutom vill man veta hur stor effekten av läkemedlet i rutinmässig användning faktiskt är. Tyvärr finns inte denna information tillgänglig i den kliniska studien utan måste skattas på något sätt.

Oavsett vilka kostnader som tas med i kostnadseffektanalysen så mäts kostnader och effekter i olika enheter. Effektmåttet väljs utifrån vad som bedöms vara den primära effektvariabeln. Med andra ord: Vilken typ av hälsoeffekt vill man i första hand uppnå med behandlingen? Tack vare att resultatmättet i kostnadseffektanalysen är förhållandevis lätt att förstå och att det bygger på samma variabler som finns i de medicinska utvärderingarna har kostnadseffektanalysen varit dominerande. I en översiktsartikel av Tengs och medarbetare [5] fanns det redan i mitten av 1990-talet över 500 kostnadseffektanalyser av preventiva program. Eftersom sådana program primärt syftar till att påverka överlevnaden tillåter det använda resultatmättet kostnad per vunnet levnadsår resultatjämförelser mellan olika typer av program.

Så snart flera effekter uppnås med behandlingen eller effekten inte är jämförbar med andra behandlingar får man dock svårigheter att tolka resultatet i den ekonomiska utvärderingen. Det kanske mest betydelsefulla bidraget till hälsoekonomin under de senaste tjugo åren är kostnadsnyttoanalys (cost-utility-analysis). I en sådan tas hänsyn inte bara till mortalitet utan även till förändring av morbiditet. En sammanvägning av eventuella förbättringar i livslängd och livskvalitet görs i måttet kvalitetsjusterade levnadsår (quality adjusted life years, QALYs). Mycket förenklat innebär det att patienters livskvalitet kartläggs och värderas med någon typ av index på en skala från 0 till 1. Noll motsvarar på denna skala död medan 1 motsvarar tillstånd med full hälsa. Om en patientgrupp i medeltal vinner två år tack vare en specifik behandling och under dessa två år har en hälsa som i genomsnitt motsvarar 0,7 så skapar behandlingen 1,4 kvalitetsjusterade levnadsår, det vill säga motsvarande 1,4 år med full hälsa.

Begreppet kvalitetsjusterade levnadsår introducerades 1964 av Herbert Klarman i en studie om kronisk njursvikt. I studien åsattes olika kvalitetsvikter på vunna år av forskarna själva. Under årens lopp har olika metoder utvecklats för att mäta preferenser för olika hälsotillstånd och för beräkning av kvalitetsvikter. Standard-Gamble-metoden innebär att försökspersoner ställs inför hypotetiska val mellan att befinna sig i ett tillstånd med nedsatt hälsa och genomgå en behandling som leder till bot samtidigt som behandlingen är förenad med en risk för en dödlig komplikation. Risken varierar tills intervjuaren tillsammans med försökspersonen funnit den risknivå där försökspersonen är indifferent inför de båda alternativen. Med hjälp av denna information kan man beräkna ett nyttovärde på det beskrivna hälsotillståndet. Nackdelen med Standard-Gamble-metoden, och en liknande metod som kallas Time-Trade-Off, är att de är tidskrävande och opraktiska att använda i kliniska studier. Istället används ofta särskilda frågeformulär som patienterna själva får fylla i. De mest vanligt använda instrumenten är: Quality of Well Being (QWB), Health Utilities Index (HUI), EQ-5D [6].

I likhet med alla andra situationer inom medicinen som bygger på värderingar förekommer givetvis individuella skillnader bland patienterna i hur de värderar sin hälsa. Dessutom finns det mättekniska problem som gör att den här typen av beräkningar måste tolkas med en viss försiktighet. Innan man blir alltför kritisk är det viktigt att fundera över hur de alternativa beslutsunderlagen ser ut. En fördel med kostnadsnyttoanalysen är att den ger ett mått, kronor per vunnet kvali-

tetsjusterat levnadsår, som gör det möjligt att jämföra utfallet mellan helt olika typer av behandlingar (livräddande och symtomlindrande) för olika sjukdomsgrupper.

Avslutningsvis några ord om kostnadsintäktanalys (cost-benefit-analysis), som är den analysmetod som är bäst teoretiskt förankrad. I korthet innebär den att både kostnader och effekt värderas i samma enhet, det vill säga kronor i Sverige. När hälsoeffekter skall värderas i pengar erbjuder detta en rad metodproblem. Det finns visserligen undersökningar av betalningsviljan för exempelvis att minska risken för skador i trafiken, som visar att värdet per vunnet levnadsår skattas till 440 000 kronor [3]. Vid motsvarande mätningar av betalningsviljan för förbättringar av olika hälsotillstånd i sjukvården missförstås dock ofta den typen av frågor. Även om metoder för att mäta betalningsvilja är under utveckling så finns än så länge få exempel på kostnadsintäktanalyser av hälso- och sjukvårdsinsatser.



Känslighetsanalys

Läsare av hälsoekonomiska utvärderingar finner inte minst på läkemedelsområdet att flertalet av dessa är så kallade modellstudier. Det innebär att författarna av dessa studier i varierande utsträckning bygger sina beräkningar på antaganden. Ibland förekommer formella simuleringar av medicinska utfall och kostnader, andra gånger har behandlingsresultat från en eller flera publicerade studier använts. En viktig del av en hälsoekonomisk studie är känslighetsanalysen. Den innebär att man undersöker hur robusta resultaten i studien är genom att pröva alternativa värden på viktiga variabler. Ofta prövar man med extrema värden från egna data eller med uppgifter från litteraturen.

Den naturvetenskapligt skolade kliniska forskaren finner det ibland märkligt att man kan basera vetenskapliga studier på antaganden. Hur kan man över huvud taget uttala sig om kostnadseffektivitet när säkra data om klinisk effekt saknas? Syftet med utvärderingar är att förbättra underlaget för beslut. Då blir frågan om resultatets relevans helt avgörande. Det problem i form av osäkra data som hälsoekonomen konfronteras med, när han eller hon skall göra en hälsoekonomisk beräkning av till exempel ett nytt läkemedel, delar han eller hon med alla andra. Ofta tvingas man fatta beslut om användning av nya medicinska metoder under osäkerhet långt innan evidensen om patientnyttan är säkerställd. I modellanalyser som inbegriper simuleringar av något slag görs beräkningar utifrån explicita antaganden.

Modellstudier är inte substitut till välgjorda kliniska studier, utan ett sätt att systematiskt sammanställa bästa tillgängliga information för att belysa ett beslutsproblem i avvaktan på säkrare data. En brist är dock att företagen och andra finansierare av hälsoekonomiska studier ofta nöjer sig med modellanalysen och på så sätt struntar i den viktiga uppgiften att verifiera/falsifiera modellen. Ett annat problem är att modellstudier i regel görs på uppdrag av eller i samarbete med ett företag med ekonomiska intressen i den produkt som studeras. Det är naturligt att så är fallet eftersom företagen har bäst kunskap om nya produkter, men företagsstudier behöver i högre grad kompletteras med oberoende analyser.

Utveckling av hälsoekonomisk metodik

I hälsoekonomisk litteratur diskuteras i ökad utsträckning hälsans fördelning i befolkningen samt vilka andra faktorer som bör beaktas vid prioriteringsutredningen av hälso- och sjukvård. Enligt den svenska prioriteringsutredningen innebär behovsprincipens överordnade natur att kostnadseffektivitetsprincipen bör tillämpas endast vid jämförelser av behandlingar av

samma sjukdom. Utredningens slutsats är att patienter med svårare sjukdomar skall gå före patienter med lindrigare sjukdomar, även om åtgärden kostar väsentligen mer är rimligt. Frågan är dock hur mycket mer det kan kosta innan det får oacceptabla konsekvenser för de grupper som därigenom måste avstå från vård? Hur skall man avgöra olika grader av behov? Hur prioriterar samhället relativt kostnadseffektiva förebyggande insatser för att sluta röka mot insatser för personer med akut ont i ryggen? I det ena fallet rör det sig om personer som i regel inte känner sig särskilt sjuka, men har en viss riskökning för framtida allvarliga händelser, medan de med ryggsmärtor är för stunden svårare sjuka medan tillgången till effektiva behandlingsmetoder är begränsad.

Ett annat exempel är läkemedlet Cerezym som används vid behandling av en svår ämnesomsättningssjukdom. Behandlingen kostar för en patient omkring två miljoner kronor per år. Kostnaden per vunnet levnadsår är många gånger högre än det som samhället tycks vara villigt att betala för att vinna levnadsår genom att minska förekomsten av rökrelaterade sjukdomar. Med andra ord så har vi en mycket hög betalningsvilja när det gäller vård av patienter med livshotande tillstånd och det finns en effektiv behandling att erbjuda, till skillnad från insatser som gäller lindrigare tillstånd men framförallt vid förebyggande av allvarliga händelser som ligger långt fram i tiden. Den här typen av skilda ställningstaganden behöver inte stå i konflikt med att ta hänsyn till kostnadseffektivitet. Frågan är hur olika värderingar på ett systematiskt sätt skall byggas in i beslut rörande fördelning av tillgängliga resurser. En utmaning för hälsoekonomer är att utveckla någon typ av »jämlikhetsvikt« som kan kombineras med kvalitetsjusterade levnadsår [7]. En sådan strävan kommer inte att leda till att beslut blir värderingsfria eller okontroversiella, men förhoppningsvis kan en sådan utveckling bidra till att tydliggöra problemet och bidra till mer rättvisa och ändamålsenliga prioriteringar.

Som framgått tidigare innehåller hälsoekonomiska analyser osäkerhet i flera led. Man brukar skilja på osäkerhet relaterad till insamlade data och osäkerhet relaterad till det sätt analysen utförts på. När det gäller osäkerheten beträffande data rör det sig i regel om en naturlig variation i populationer som kan minskas på sedvanligt sätt genom att öka studiepopulationens storlek eller genom att beakta osäkerheten med hjälp av ett konfidensintervall. Att beräkna konfidensintervall för en kvot är dock mer komplicerat än att beräkna dem för enstaka medelvärden. Den hälsoekonomiska litteraturen erbjuder numera flera metoder för att ange osäkerhetsintervall vid beräknade kostnadseffektkvoter [8]. Det kommer dock säkerligen att ta flera år innan dessa slagit igenom i praktiken.

Ett tredje område som uppmärksammas är svårigheterna att jämföra kliniska och ekonomiska data från multicenterstudier med data från olika länder. Kostnadseffektkvoterna kan då variera i olika delmaterial, på grund av olika utfall på effektsidan, praxisskillnader och olika prisnivåer. Ett sätt att hantera detta kan vara att beräkna länderspecifika kostnadseffektkvoter och analysera dem separat innan en beräkning görs för hela studiepopulationen [9]. Inom detta område finns det många metodproblem som väntar på sin lösning.

Ett fjärde metodområde inom hälsoekonomi gäller sammanvägning av resultat från flera studier. Metaanalyser av hälsoekonomiska studier är ovanliga av flera skäl. Ett är att det hittills sällan funnits flera studier som analyserat samma frågeställning, ett annat att det i regel är svårare än med kliniska data att jämföra resultaten från flera ekonomiska utvärderingar. Det är inte heller många år sedan Cochrane Collaboration och andra liknande organisationer presenterade rikt-

linjer för systematisk granskning av kliniska studier. När det gäller granskning och sammanvägning av hälsoekonomiska studier är metodutvecklingen fortfarande i sin linda. Det pågår utveckling på flera håll, och inom några år har vi säkert fått bättre verktyg för att göra adekvata sammanställningar av den växande hälsoekonomiska litteraturen. Under tiden rekommenderas en bra artikel av Drummond och Jefferson som handlar om riktlinjer till författare och vetenskapliga granskare av studier som insänts till BMJ [10].

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Jansson M, Levin LÅ, Swahn E. Cost-effectiveness of an invasive strategy in unstable coronary artery disease. *European Heart Journal* 2002;23:31-40.
2. Gold M, Siegel J, Russel L, Weinstein M, editors. Cost-effectiveness in health and medicine. Oxford: Oxford University Press; 1996.
3. Ekman M, Zethraeus N, Dahlström U, Höglund C. Kostnadseffektivt att behandla kronisk hjärtsvikt med bisoprolol. *Läkartidningen* 2002;99:646-50.
4. SBU. Trombocythämmande läkemedel – klopidogrel (Plavix). Available from: URL: <http://www.sbu.se>
5. Tengs TO, Adams ME, Pliskin JS, Safran DG, Siegel JE, Weinstein MC, et al. Five-hundred life-saving interventions and their cost-effectiveness. *Risk Anal* 1995;15(3):369-90.
6. Henriksson M, Carlsson P. Att mäta hälsorelaterad livskvalitet – en beskrivning av instrumentet EQ-5D. Linköping: Linköpings universitet; 2002. CMT-rapport 2002:1.
7. Lindholm L. Kostnadseffektivitetsanalys som redskap för prioriteringar. Bilaga 5. Prioriteringar i vården, Prioriteringsdelegationen, Socialdepartementet. SOU 2001:8.
8. Briggs A. Handling uncertainty in economic evaluation and presenting the results. In: Drummond M, McGuire A, editors. *Economic evaluation in health care. Merging theory with practice*. Oxford: Oxford University Press; 2001.
9. Drummond M, Pang F. Transferability of economic evaluation results. In: Drummond M, McGuire A, editors. *Economic evaluation in health care. Merging theory with practice*. Oxford: Oxford University Press; 2001.
10. Drummond M, Jefferson T for the BMJ Working Party. Guidelines for authors and peer reviewers of economic submissions to the BMJ. *BMJ* 1996;313:275-83.

Särtryck

Läkartidningen

Alla kroppens celler reagerar på olika signalämnen i omgivningen, ämnen som styr deras fundamentala livsprocesser.

Dessa ämnen kallas kollektivt tillväxtfaktorer. En serie i *Läkartidningen* 1995 om dem speglar tendenser i dagens medicinska forskning och pekar på några tillämpningsområden.

Området är i början av en snabb utveckling och många produkter är under utprövning för klinisk användning.

Häftet omfattar 12 artiklar på sammanlagt 56 sidor + färgomslag.

Priset är 75 kronor

Tillväxtfaktorer



Beställer härmed.....ex
av "Tillväxtfaktorer"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till **Läkartidningen**
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Hans Liedholm, docent, överläkare (hans.liedholm@smi.mas.lu.se)

Agneta Björck Linné, med dr, apotekare; båda vid Universitetssjukhuset MAS, Malmö (agneta.bjork-linne@smi.mas.lu.se)

Att ranka läkemedels effekt – en »klassfråga«?

Det är svårt att bedöma om två behandlingar är likvärdiga

II Med vilken trovärdighet kan en expertgrupp i en läkemedelskommitté t ex föreslå att fluvastatin skall rekommenderas i stället för pravastatin vid lipidrubbnings, eller att bisoprolol skall rekommenderas i stället för metoprolol vid hjärtsvikt, utan att det finns en enda publicerad jämförande långtidsstudie med dessa läkemedel med kliniskt väsentliga effektvariabler som effektmått? Båda exemplen kan vara uttryck aningen för en brist på förståelse av evidensbaserad medicin (EBM) eller för en önskan att spara pengar åt landstingen genom rekommendationslistorna. Det är förstås underförstått att expertgruppen uppfattat att alternativen är utbytbara. De två alternativen i varje grupp befinner sig ju i samma »klass«?

Under de senaste åren har läkemedels utbytbarhet debatterats i några ledande medicinska tidskrifter [1-4]. Frågan har gällt förekomsten av »klasseffekt« när det gäller läkemedel. Kan man över huvud taget placera nya, likartade substanser i samma klass som äldre medel om de inte undersökts på likartat sätt i omfattande kliniska studier?

Hur mycket vi behöver ta strikt hänsyn till närvaro av eventuell klasseffekt är avhängigt av indikationerna och läkemedlens bevisade effekter. För symtombehandling – t ex behandling av smärta, klåda, hosta – är dokumentationskraven på en klasseffekt relativt sett mindre. Vid tung sjukdomsbehandling – t ex kliniska effekter på mortalitet, invaliditet, inläggningssfrekvens på sjukhus, operationsbehov, eller där den eftersträvade effekten (att förhindra kardiovaskulär död eller insjuknande) inte går att i realtid utvärdera hos den enskilde patienten – är dokumentationskravet mycket stort.

Vad är läkemedelsklass?

Begreppet läkemedelsklass saknar en tydlig definition. När vi diskuterar vilka egenskaper ett nytt läkemedel har vill produktens förespråkare gärna förmedla hur speciellt eller unikt medlet är. Det sägs kanske rentav tillhöra »en ny klass« eller »en ny generation«. Om det redan finns, eller snart tillkommer, fler läkemedel som placeras i samma ATC-grupp (Anatomical Therapeutic Chemical Classification) ökar benägenheten att föra samman substanserna till en klass.

Inom WHO, och därmed i många länder, används ATC-systemet för att klassificera läkemedel. Förkortningen står för att det är ett anatomiskt, terapeutiskt och kemiskt klassifika-

SAMMANFATTAT

Var skeptisk mot försök att generalisera resultat inom samma läkemedelsgrupp. Begreppet »klass« är vagt definierat.

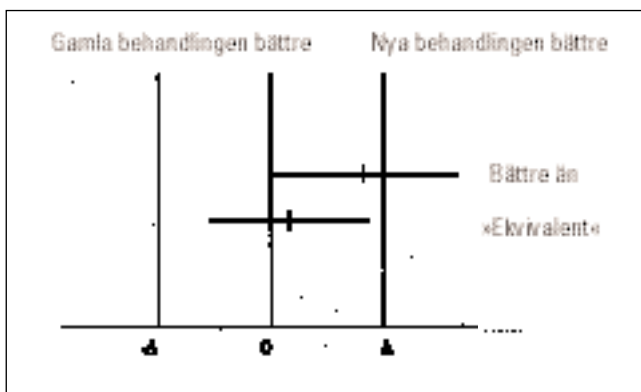
Om en studie inte visar någon skillnad mellan två terapier, bedöm om ekvivalens var det primära målet med studien eller om slutsatsen baseras på en efterkonstruktion som inte visat överlägsenhet för den nya åtgärden.

Om syftet med studien var att visa ekvivalens, bestäm om definitionen av ekvivalens eller den minsta, kliniskt värdefulla skillnaden var rimlig och om analysen stöder slutsatserna.

Om analysen av resultaten ger stöd för ekvivalens, bestäm om andra faktorer är tillräckligt övertygande för en rekommendation om användning av den nya behandlingen i stället för den traditionella skall kunna ges.

Evidensbaserad medicin

tionssystem. Anatomi syftar på det anatomiska terapiområdet (t ex hjärta), terapi syftar på terapiområdet (t ex att sänka blodtryck), kemi syftar på kemisk gruppering (t ex tiazid). Om man betraktar uppställningen i ATC-registret i Fass över de nio kalciumantagonisterna (ATC-kod Co8) framgår att det finns skillnader mellan och inom substanserna i gruppen. I den största gruppen, dihydropyridinerna, tillkommer uppgifter om beredningsformer (tablett, filmdragerad tablett, depåtablett, depåkapsel och lösning till infusionskoncentrat), vilka åtföljs av variationer i farmakokinetiska egenskaper och effekt. ATC-systemet hjälper oss inte att klara ut problemet med klasseffekt så som en kliniker ser det. Hos FDA (Food



Figur 1. Schematisk bild av olika behandlingseffekter. De vertikala strecken (Δ) representerar gränser ovan och under en effekt (0) visad för ett jämförelsepreparat. Gränserna sätts utifrån någon rimlighetsnivå: vad som är kliniskt intressant att överträffa eller understiga. De horisontella strecken representerar konfidensintervall för en behandlingseffekt i förhållande till jämförelsebehandlingen. I det övre området överstiger effekten 0, medan effekten i det undre området ligger mellan de båda gränserna (fritt efter Hans Melander, Läkemedelsverket, Riksstämman 2000 och Massel [4]).

II Fakta

Checklista vid bedömning av undersökningar som hävdar »likhet«

- Var studien designad för att påvisa ekvivalens?
- Är jämförelseinterventionen effektiv? Medför den nya behandlingen någon fördel framför den gamla terapin?
- Beskrevs den minsta, kliniskt betydelsefulla skillnaden redan på planeringsstadiet av studien, före den primära analysen?
- Kan någon form av bias förklara resultatet av »icke-skillnad«?
- Redovisade man alla patienter som ingick i studien?
- Är det möjligt att grupperna skiljer sig mindre än den minsta betydelsefulla skillnaden?
- Gavs behandlingarna precis på samma sätt som vid den vanliga behandlingen av tillståndet?
- Är fördelarna med den »ekvivalenta« behandlingen värd den potentiella skadan och kostnaden?

and Drug Administration) kommer definitionen närmare: »All products within a class are assumed to be closely related in chemical structure, pharmacology, therapeutic activity, and adverse reactions« [1].

En läkemedelsklass kan bestå av en grupp av läkemedel som har likartad

- kemisk struktur (t ex dihydropyridinringen hos dihydropyridinkalciumantagonister),
- verkningsmekanism (kalciumantagonister blockerar spänningsberoende kalciumkanaler på cellmembran),
- farmakologisk effekt [2].

I det sista kriteriet bör både klinisk effekt och bieffekt ingå. Som vi skall se senare kan det vara svårt att hålla ihop klassen om inte kriterierna uppfylls.

När kan vi räkna med att en klasseffekt föreligger?

Den absoluta behandlingseffekten varierar mellan patientgrupper då de studeras med olika slags terapi. Vid jämförelser av resultaten görs jämförelsen bäst mellan proportionella ef-

fekter, t ex relativ riskreduktion (RRR) eller oddskvot/oddsratio (OR). Om läkemedel med samma verkningsmekanism ger RRR eller OR av samma storleksordning och riktning kan de betraktas ha en klasseffekt. Så visade t ex en jämförelse av 32 studier av ACE-hämmare på 7 105 patienter med hjärtsvikt likartade effekter [5]. Effekten, mätt som total mortalitet, varierade en hel del, men konfidensintervallen för effekt för de olika ACE-hämmarna överlappade varandra och resultaten för de olika medlen var inte heterogena (pekade inte i olika riktningar). Det behöver knappast sägas att en kritisk punkt för en korrekt bedömning därvid är tillgång till all väsentlig dokumentation, även opublicerad sådan. För att ta reda på hur olika behandlingar förhåller sig till varandra behöver de undersökas i kontrollerade försök, dvs kliniska prövningar.

Mål med kliniska prövningar

Avsikten med en klinisk prövning skall givetvis alltid definieras före studiestart. Man kan hävda att det finns tre typer av kliniska prövningar: att försöka visa överlägsenhet, att försöka visa ekvivalens eller att testa att ett medel inte är sämre än ett jämförelsepreparat. Mer om detta står att läsa i en översiktsartikel [4]. De statistiska underlagen för bedömningarna framgår av tillgängliga dokument hos EMEA, den europeiska registreringsmyndigheten [6, 7]. Dokumenten kan läsas på EMEAs webbplats <http://www.eudra.org/emea.html>.

Överlägsenhet. De flesta kliniska prövningar konstrueras för att påvisa överlägsenheten, »superiority study«, av en intervention (t ex ett läkemedel) över en annan, och dimensioneras för detta.

Ekvivalens. Mindre vanligt är att försöka visa att en liten skillnad är tillräckligt liten för att behandlingarna skall betraktas som utbytbara eller kliniskt ekvivalenta. En sådan studie brukar benämnas ekvivalensstudie, »equivalence study«. Studien behöver då ofta vara stor, eftersom det med bibehållen styrka statistiskt sett kräver fler patienter att påvisa en mindre skillnad. Om ekvivalens skall sägas råda måste i en ekvivalensstudie denna skillnad beskrivas, vara kliniskt realistisk och inte överskridas. En metodologisk lärorik artikel och medföljande kommentar till en ekvivalensstudie av HIV-läkemedel har publicerats i JAMA [8, 9].

Ingen skillnad är inte detsamma som likhet! Det är mycket viktigt att noggrant göra åtskillnad mellan en studie som avsett visa ekvivalens och en studie som avsett visa överlägsenhet men som inte lyckats göra det och till slut hävdar »likhet«. Efter ett sådant resultat försöker ibland författarna (och sponsorererna) hävda att behandlingarna är lika bra.

Ett färskt exempel är marknadsföringen av NORDIL-studien, som avsåg att visa överlägsenhet av diltiazembaserad terapi över traditionell terapi [10]. När detta inte blev fallet har man i stället hävdats att diltiazembaserad terapi av högt blodtryck är lika bra som behandling baserad på traditionell terapi. En liknande felaktig slutsats har dragits av resultatet av en nifedipinbaserad blodtrycksstudie, INSIGHT [11]. Studierna har kommenterats i Evidence Based Medicine [12].

Icke sämre än. I en tredje typ av studie (som anges öka i förekomst) försöker man fastställa t ex att ett läkemedel i varje fall inte är sämre än det man jämför med, »non-inferiority study«. Studier med denna frågeställning ställer mycket stora krav på design och tolkning och diskuteras inte ytterligare här. En schematisk bild över de två första typerna av studier visas i Figur 1.

Vad krävs för att klasstillhörighet skall kunna hävdas?

För att bedöma evidensen i dokumentationen av läkemedelseffekter bör man använda olika kriterier, som kan anges gra-

Tabell I. Evidensnivåer för effekt. RCT= randomised controlled trial. Direkt jämförelse = jämförelse mellan två preparat i samma studie. Indirekt jämförelse = jämförelse mellan två preparat där dessa undersökts var för sig i olika studier.

Nivå	Jämförelse	Patienter	Effekter	Kommentarer
1	RCT direkt jämförelse	Identiska	Kliniskt viktiga	Kräver dolt randomiseringsschema, jfr [14], fullständig uppföljning, dubbelblindning och »vettig« effektvärdering
2	RCT direkt jämförelse	Identiska	Valida surrogatmått	Som nivå 1 plus validerat surrogatmått
2	Indirekt jämförelse med placebogrupp från olika RCT	Likartade eller olika beträffande sjukdomsgrad eller -risk	Kliniskt viktiga eller valida surrogatmått	Som nivå 1 plus angivande av eventuella skillnader i metodologisk kvalitet beträffande händelser, compliance och risk vid start
3	Subgruppsanalyser från indirekta jämförelser av olika RCT med placebogrupp	Likartade eller olika beträffande sjukdomsgrad eller -risk	Kliniskt viktiga eller valida surrogatmått	Som nivå 1 men risker med multipla jämförelser, post hoc »data dredging«, undergrupper med låg »power« och missklassifikationer till subgrupper
3	Indirekt jämförelse med placebogrupp; från olika RCT	Likartade eller olika beträffande sjukdomsgrad eller -risk	Ovaliderade surrogatmått	Surrogathändelser fångar kanske inte upp alla goda eller dåliga effekter av behandlingen
4	Indirekta jämförelser av icke randomiserade studier	Likartade eller olika beträffande sjukdomsgrad eller -risk	Kliniskt viktiga	Risker med »Confounding by indication«, »compliance« eller tid; okända eller omätta confounders, mättningsfel; data hämtade från begränsade databaser eller användning av kodningssystem olämpliga för forskning.

derade i olika nivåer (Tabell I). Uppgifterna i tabellen är en bearbetning av beskrivningar från [2] och [13].

På evidensnivå 1 finns de studier som undersöker kliniskt betydelsefulla händelser och som utgör primär effektvariabel. Här finns studier (randomiserade kontrollerade studier) som är »head-to-head«-jämförelser, dvs jämförelser av minst två läkemedel inom samma studie. Vilka speciella händelser som skall väljas beror på det tillstånd och den undersökningstid som undersöks. För statiner, som primärt används för att sänka lipidnivåer för att förhindra eller behandla aterosklerotiska sjukdomar, kan de kliniskt viktiga händelserna bestå av t ex total dödlighet oavsett orsak (»all-cause mortality«), hjärtinfarkt och slaganfall. Studier med lägre evidensnivå (nivå 2–4) har ofta många metodologiska problem, som gör att resultaten från sådana undersökningar är svaga.

Faktarutan ger en lista över punkter som är till hjälp då man skall bedöma en undersökning som hävdar »likhet« [4]. De två sista punkterna är värdefulla i praktisk klinik och bör stimulera till en konstruktiv diskussion kolleger emellan.

Effekten av den aktiva terapi man jämför med bör tidigare ha visats i (minst) en stor randomiserad kontrollerad studie med låg risk för typ I- och typ II-fel eller i en välunderbyggd metaanalys [15]. Om resultaten av den nya behandlingen inte visar på någon överlägsenhet bör åtminstone någon fördel visas beträffande minskad kostnad, kortare sjukhusvård, färre inläggningar på sjukhus, färre biverkningar, mindre besvär för patienterna, lättare hantering för sjuksköterskor etc för att göra den nya terapin intressant.

Bedömdes alla kliniskt viktiga händelser? Var storleksordningen av skillnaden ett rimligt mått på ekvivalens, och kunde detta mätas med god precision? Var inklusionskriterierna av patienterna tillräckligt väl beskrivna? Var randomiseringen invändningsfri? Var grupperna lika från start? Bibehölls maskeringen? Behandlades grupperna lika?

Det finns stor risk för att data påverkas när en studie sträcker sig över en lång tidsperiod. Samtiden inför successivt nya rutiner och behandlingar och en »kontaminering« av den nya terapin kan smyga sig in i kontrollgruppen. Risk finns också för annan påverkan när studier genomförs i flera länder med

skiftande rekryteringsgrad och med olika biologiskt utfall av det studerade tillståndet i olika miljöer.

Analyseras både efter principerna »intention to treat« och »on treatment«? Var uppföljningen fullständig? Var tidsramen för uppföljning tillräcklig för att dra slutsatsen att ekvivalens förelåg? Studier med stort bortfall – exempelvis studier av kroniskt obstruktiv lungsjukdom, alkoholism, demens, HIV-infektion – medför automatiskt att validiteten sjunker. Absolut ekvivalens kan aldrig bevisas, men det bästa man kan åstadkomma är att visa att det är osannolikt att den sanna skillnaden (inom 90 eller 95 procents säkerhet) är större än den minsta betydelsefulla skillnad man är intresserad av.

Gavs behandlingarna på samma sätt som vid vanlig behandling av tillståndet? Detta är en ofta förbisedd fråga. Grundar sig studieresultatet på en ovanligt låg eller ovanligt hög dos av läkemedlen? Användes den vanliga beredningsformen, eller var den kanske inte ens angiven? Användes hjälpmedel för administreringen som inte finns i vårt land, exempelvis tekniska hjälpmedel för inhalation? Kan resultaten överföras till den egna patientgruppen?

Exemplet klopidogrel–ASA

Diskussioner kring terapier som är definierade som ekvivalenta förekommer knappast i praxis. Användningen av klopidogrel, som i praxis uppfattas vara likvärdigt ASA (men som vetenskapligt inte har den bakgrunden), är dock ett exempel på hur en debatt skulle kunna läggas upp. Om preparaten bedömdes ekvivalenta, hur mycket är man beredd att spendera? Enligt SBU Alerts bedömning över klopidogrel (<http://alert.sbu.se/index.asp?lang=SWE>) skulle en övergång från ASA (1998) ge en merkostnad på 6 000 kronor per patient och år. Om klopidogrel i stället användes vid 10 procent av alla behandlingsdagar med ASA skulle detta innebära att läkemedelskostnaderna i 1998 års siffror ökade med uppskattningsvis 250 miljoner kronor.

Klasseffekter

Blodtryckssänkande medel. Frågan om klasseffekt kan uppfattas handla inte bara om substanser inom samma farmako-

Tabell II. Frånvaro av bevis eller bristande bevis för effekt för bedömning av eventuell klasseffekt.

Läkemedelsgrupper	Användning	Referens	Klasseffekt?
ACE-hämmare	Tillägg vid hjärtsvikt	3	Evidens för effekt på »hard endpoints« föreligger ej för alla
All-antagonister	Blodtryckssänkning	42-44	Evidens för effekt på »hard endpoints« otillräcklig. Få jämförande studier
Atenolol Bucindolol	Tilläggsbehandling vid hjärtsvikt	34 35	Ej välundersökt Försök finna evidens misslyckad
Atenolol	Profylax efter hjärtinfarkt		Ej välundersökt
Perorala diabetesmedel	Diabetes		Evidens för effekt på »hard endpoints« föreligger inte för alla. Jämförelser ej välundersökta
Neuroleptika, atypiska neuroleptika	Schizofreni		Ännu få jämförande studier där effekt och bieffekter jämförts
NSAID – »oselektiva« COX-1 sparande	Smärta, inflammation		Ännu få jämförande studier där effekt och bieffekter jämförts
Kolinesterashämmare	Alzheimers sjukdom		Evidens för effekt på »hard endpoints« föreligger ej; inga jämförande studier
Bisfosfonater	Frakturförebyggande		Evidens för effekt på »hard endpoints« föreligger ej för alla. Jämförande studier saknas
Tiludronat, ibandronat		45, Ö Ljunggren, pers medd, 2001	Försök finna evidens misslyckad
Tiazider, betablockerare, kalciumantagonister ACE- hämmare, All-anta- gonister, alfablockerare m fl	Blodtryckssänkning. Skydd mot kardiovaskulära händelser	17-19, 23-26	Evidens för effekt på »hard endpoints« föreligger ej för alla. Brist på jämförande studier (evidensnivå 1)
Statiner	Kolesterolsänkning	2, 46-48	Evidens för effekt på »hard endpoints« föreligger ej för alla. Få jämförande studier
Simvastatin, fluvastatin	Kolesterolsänkning – förebygga stroke	49, 50	Olikheter i evidens
GP IIb-/IIIa-antagonister	Trombocythämning	51	Olikheter i evidens
Trombocythämmare, tiklopidin klopidogrel		52–55	Brist på jämförande studier. Varierande dokumentation

logiska grupp utan också om likartad terapieffekt hos olika grupper av läkemedel. Ett exempel på det senare är en ofta ställd fråga om olika blodtryckssänkande medels effekt: bäst eller lika bra? Om ett läkemedel sänker blodtrycket och har bevisat god effekt också på mortalitet och morbiditet kan man väl anta att andra substanser som också sänker blodtrycket delar dessa effekter? Vid närmare eftertanke är detta inte givet. En inblick i fysiologiska regleringar för normalt och förhöjt blodtryck bör få oss att tveka om huruvida resonemanget är korrekt [16].

Kritiska genomgångar av dokumentationen över blodtryckssänkande läkemedel har publicerats [17-19]. I välkända referensverk åtskiljer man inte läkemedel inom klasserna [20-22], med undantag för kortverkande mot långverkande

dihydropyridiner. Substanser inom samma grupp av äldre läkemedel, t ex diuretika, bedöms alla ha samma egenskaper och samma kliniska effekt. Detta är kanske korrekt, men i den grupp som innehåller flera substanser med kemiskt snarlik struktur har de kliniska långtidsresultaten och egenskaperna hos olika tiazider inte undersökts mot varandra »head to head« i randomiserade, kontrollerade studier. De undersökningar som skulle erfordras för detta kommer förstås inte att genomföras. Motsvarande gäller många andra medel.

Två omfattande metaanalyser av skillnader mellan olika antihypertensiva läkemedel publicerades samtidigt häromåret [23, 24]. Särskild uppmärksamhet ägnades dokumentationen av kalciumantagonister. I den ena analysen visades att kalciumantagonister gav sämre skydd mot infarkt, hjärtsvikt

Tabell III. Evidensnivåer för biverkningsdata (efter Bandolier [13]).

Nivå	Typ av studie	Fördelar	Kommentarer
1	RCT	Den enda design som tillåter upptäckt av biverkningar när biverkningen är liknande den som behandlingen skall försöka förhindra	Risk för alltför låg »power« för att upptäcka biverkningar såvida inte studien designats för att göra det
2	Kohort	Prospektiv datainsamling, definierad kohort	Är kritiskt beroende av noggrannhet i uppföljning, klassificering och mätningförfarande
3	Fall-kontroll	Billig, vanligen snabb att genomföra	Selektions- och minnesbias kan ge problem, temporala förhållanden är kanske inte tydliga
4	Fas IV	Kan upptäcka ovanliga och allvarliga händelser om den görs stor nog	Saknar kontroll eller har kontroll som inte är matchad. Är kritiskt beroende av noggrannhet i uppföljning, klassifikation och mätningar
5	Fallsérie	Billig, vanligen snabb att genomföra	Ofta små, selektionsbias kan utgöra problem, saknar kontrollgrupp
6	Fallrapport(er)	Billiga, vanligen snabba att genomföra	Ofta små, selektionsbias kan utgöra problem, saknar kontrollgrupp

Tabell IV. Godtyckligt valda exempel på skadliga effekter av substanser som i flera exempel lett till dödsfall och indragning från marknaden, oavsett om viss dokumentation redan förelagat över terapeutiska egenskaper som kunnat föra till inklusion i läkemedelsklasser. Uppmärksamhet på biverkningarna har lett till omvärdering av respektive klass.

Läkemedel eller läkemedelsklass	Användning	Referens	Kommentar
Doxazosin	Blodtryckssänkning	61	Hjärtsvikt
Pronetanol, Praktolol	Betablockering, blodtryckssänkning	36, 37	Misstänkt cancerogent, hud- och slemhinnebiverkningar
Kalciumantagonister, kortverkande DHP	Blodtryckssänkning	23, 24	Kardiovaskulära skador med kortverkande substanser, särskilt nifedipin
Mibefradil	Hjärtsvikt	62, 63	Allvarliga läkemedelsinteraktioner
Cerivastatin	Kolesterolsänkning	56, 57	Muskelpåverkan, njurskada
Benoxaprofen, indometacin (osmos) isoxicam, oxyfenbutazon, fenybutazon	Smärtstillning	64, 65	Blödningar och perforationer i mag-tarmkanalen, blod- och njurpåverkan m m
Trova-, tema-, grepa-, spar, lome-, tosu-, moxi-, clina-, pefloxacin; fleroxacin, enoxacin	Infektionsbehandling	66	Påverkan på lever, hud, blod, CNS, njurar, hjärta, blodglukos
Zimelidin, grad av selektivitet av senare SSRI	Depressionsbehandling	67-70	Guillain-Barrés syndrom, blödningsbenägenhet vid ökad selektivitet
Remoxiprid, clozapin, ziprasidon	Schizofrenibehandling	71-75	Neutropeni, QT-förlängning, hjärtarytmi
Troglitazon	Blodsockersänkning	76, 77	Leverpåverkan
Terfenadin, ketokonazol, astemizol	Allergibehandling	78-80	QT-syndrom, hjärtarytmi
Takrin	Alzheimerbehandling	81	Leverpåverkan
Tiklid, Klopido-grel	Trombocythämning	82, 83	Neutropeni, trombotisk trombocytopen purpura
COX-I-sparande, NSAID, Rofecoxib	Smärta, inflammation	60	Ökade kardiovaskulära biverkningar vid samtidig minskning av slemhinneskador i mag-tarmkanalen
Aptitnedsättande medel, Fenfluramin, Dexfenfluramin	Överviktsbehandling	84, 85	Hjärtklaffspåverkan

och större kardiovaskulära händelser («number needed to harm», NNH, 110, 152 respektive 90). Däremot var siffrorna för total dödlighet eller insjuknande i slaganfall inte åtskilda. I den andra analysen bedömdes resultatet avseende insjuknandet i koronarsjukdom som marginellt sämre, medan siffrorna för insjuknande i slaganfall var signifikant bättre för kalciumantagonister än för diuretika och betablockerare («number needed to treat», NNT 172).

Kommentatorn i Evidence Based Medicine gav inte de två metaanalyserna högsta betyg, på grund av framför allt brist på data [25]. Kommentatorn var dessutom missnöjd med att olika substanser inom en grupp sammanförts till samma klass. Båda metaanalyserna ansågs dock viktiga, och klart var att vissa diuretika, betablockerare, ACE-hämmare och kalciumantagonister visat positiva effekter beträffande morbiditet och mortalitet för vissa personer med högt blodtryck. Problemet kvarstår att identifiera de individrelaterade egenskaperna. Debatten fortsätter; förra året kom ytterligare en meta-analys – med en annan tolkning [26].

Betablockerare. Man anser att det finns endast två bedömbara randomiserade studier av betablockerare vid hypertoni som jämfört effekten mot placebo och visat på skyddande effekter beträffande hjärt-kärlskador [18]. Andra studier, som ofta anförs som evidens för effektivitet av betablockerare utöver tiiazider, är inte användbara, då respektive behandlingsgrupp i så stor utsträckning fått tillägg av annan terapi [27, 28]. Så långt finns det alltså alltför få studier för att svara på frågan om klasseffekt.

Kan man inte jämföra effekter av betablockerare som givits efter hjärtinfarkt där många patienter ändå samtidigt haft indikationen hypertoni [18, 29]? I placebokontrollerade studier med en duration på mer än tre månader och mortalitet som slutmätning har resultaten skiftat. Signifikant bättre resultat har visats för de icke-selektiva preparaten propranolol och timolol, inga för selektiva preparat och endast för acebutolol bland preparat som är partiella agonister (med sk egenstimulerande effekt, ISA) [18]. Vid en genomgång av alla kontrollerade långtidsstudier (omfattande mer än sex månader) fanns stark evidens för effekt på mortalitet och morbiditet för propranolol, timolol och metoprolol [29]. Orsaken till skillnaderna mellan olika preparat är svår att tolka, då befintliga betablockerare har egenskaper som ger stor variation [30]. Beta 2-blockad, lipidlöslighet, längden på betablockeringen över dygnet, patienturval mm kan var för sig ha bidragit.

Ett exempel på den komplexa bilden är en studie av d-isomeren av sotalol. Den saknar visserligen de betablockerande egenskaper som racematet sotalol (dl-sotalol) har, men är verksam som kaliumkanalblockerare. I en studie av hjärtsvikt efter hjärtinfarkt fann man en ökad mortalitet (vanligen hjärtarytmi), vilket ledde till att studien stoppades [31].

Möjligheten att det skulle finnas ett slags klasseffekt hos betablockerare har ändå hävdats. I en epidemiologisk studie jämfördes kliniska data från patienter som skrivits ut efter en hjärtinfarkt och ordinerats atenolol, metoprolol eller propranolol med data från obehandlade kontroller [32]. Tydliga fördelar fanns av behandling med betablockerare, men då tvåårs-mortaliteten var ganska lika i de tre behandlade grupperna ansåg man att det kunde tyda på att valet av preparat inte spelat någon större roll, varför man drog slutsatsen att det fanns en klasseffekt. Den naturliga invändningen mot denna slutsats är förstås att det inte var en randomiserad studie.

Vårt behov av att uppfatta sammanhang som inte undersöks påverkas ibland av nya data. Det fanns positiv dokumentation över effekt av betablockerare vid hjärtsvikt för minst sju olika betablockerare [33]. Senare studier av karvedilol, bisoprolol och metoprolol har påvisat evidens för effekt

mot död och försämring vid administrering till patienter med systolisk hjärtsvikt. Samtidigt med en positiv studie över karvedilol vid svår hjärtsvikt kom en studie av bucindolol, som inte kunde påvisa någon positiv effekt [34, 35]. Frågan om klasseffekt komplicerades.

Då och då upptäcks preparatspecifika biverkningar som förändrar uppfattningen om klasseffekt. En av de första betablockerarna, pronetalol, försvann från marknaden på grund av en risk för att den skulle kunna vara kancerogen [36]. Pronetalol ersattes sedan av propranolol. En tidig, beta 1-selektiv betablockerare från samma tillverkare var praktolol. Efter cirka fyra år på marknaden och en miljon patientår upptäcktes ett speciellt biverkningssyndrom, ett sk okulomukokutant syndrom, som kunde bestå av psoriasisliknande utslag i handflator och på fotsulor, keratoconjunctivitis sicca, serös otit, med dövhet, skleroserande peritonit, pleurit och perikardit [37]. Upptäckten fick stor inverkan på det fortsatta kliniska och prövningsmässiga arbetet.

ACE-hämmare. Det finns inga studier av ACE-hämmare som visar mortalitetsminskning hos patienter med okomplicerad hypertoni [3]. Därför kan de inte anses jämbördiga med t ex tiiazider. Vid hjärtsvikt efter hjärtinfarkt är bristen på dokumentation inte föraktlig från klassynpunkt, och endast de substanser som visat effekt på mortalitet och morbiditet kan rekommenderas [3]. På andra variabler – som effekt på njurfunktion hos diabetiker och hos patienter med annan njursjukdom – har kontrollerade studier visat gynnsamma effekter. Hela gruppen ACE-hämmare har dock inte denna dokumentation.

Ett annat exempel där jämförande studier av andra ACE-hämmare saknas är den sk HOPE-studien, där behandling med ramipril jämfördes med placebo och visat positiva effekter hos patienter med olika typer av risk för kardiovaskulär sjukdom [38]. Kommentatorn i ledartikeln ansåg inte att effekten var bedömbär som en klasseffekt, och kravet på jämförande studier var stort [39]. Inte oväntat hävdades snart ändå motsatsen [40]. Med tanke på kända skillnader mellan substanserna i gruppen är det lätt att förstå att problemet inte är löst [3, 41]. I Tabell II illustreras frånvaron av bevis eller bristande bevis för effekt.

Biverkningar och indragningar

En viktig fråga när det gäller klasseffekt är således effektmässiga likheter. En annan fråga rör biverkningar. Även om inte regelrätta undersökningar finns som underlag accepteras »likhet« snabbt i den kliniska vardagen, lika snabbt kan preparatspecifika skillnader uppvärderas i marknadsföringen. Information om biverkningar spräcker mönstret, särskilt om substanser då och då måste dras in. Vid närmare analys framkommer kanske att även andra substanser i gruppen har sådana biverkningar. Vid diskussionerna förra året om dödliga muskel- och njurbiverkningar av cerivastatin framkom t ex att alla statiner – om än i mindre omfattning och kanske under andra förhållanden – hade rapporterats ha sådana biverkningar [56, 57].

Biverkningar av läkemedel kommer förr eller senare att påverka balansen. En oriktig redovisning rubbar grunden för rättvisa jämförelser. I en systematisk genomgång av biverkningar i 192 randomiserade kontrollerade studier fann man att det förelåg en besvärande underrapportering [58]. Bara i knappt hälften av studierna hade man rapporterat de specifika skälen till att behandlingen avbrutits på grund av biverkningar. Sena exempel på oriktig redovisning av effekter och biverkningar som påverkar helhetsbilden är diskussionen på FDA om de COX 1-sparande medlen celecoxib och rofecoxib [59, 60].

Det är lämpligt att försöka bedöma biverkningsproblem i

förhållande till evidensnivån. Förenklat illustreras detta i Tabell III. Exempel på substanser som mer eller mindre oväntat förknippats med biverkningar ges i Tabell IV.

Konsekvenser och oväntade resultat

Ett välment byte från simvastatin till fluvastatin gjordes på Nya Zeeland eftersom de båda substanserna ansågs likvärdiga [86]. Vid kontroll av effekten fann man överraskande att lipidnivåerna hade stigit, liksom antalet vaskulära händelser.

Ett annat oväntat resultat belystes i en kanadensisk observationell studie. ACE-hämmarna kaptopril, enalapril och lisinopril hade fått ett »referenspris« för att de bedömts likvärdiga. De patienter som initialt fått dessa läkemedel jämfördes. Efter fyra års uppföljning visade det sig att patienter som behandlats med kaptopril använde mera sjukvårdsresurser än de andra, en oväntad effekt av behandlingen [87].

En faktor som kan förändra omdömet om eventuell klasseffekt är då något annat än själva läkemedlet påverkar resultatet i en studie. I en kontrollerad studie över förmågan hos en äldre befolkning att lära sig använda det ordinarie hjälpmedlet för inhalation av influensaläkemedlet zanamivir befinns detta vara både krångligt att hantera och att memorera [88].

Diskussion

Omfattande diskussioner förs avseende kostnadsmässigt viktiga terapiområden som diabetes, ulkus och refluxesofagit, »hormone replacement therapy« (HRT), demens, osteoporos, schizofreni, trombocythämning, smärta och inflammation. Argument hämtas från kliniska prövningar, tematiska framställningar, specialister inom områdena, riktlinjer, marknadsföring av läkemedelsindustrin och deras budbärare inom den medicinska professionen. Det finns många uppfattningar om likheter och olikheter.

Vid vår genomgång kan vi se att frågan om läkemedelsklasser och klasseffekter mycket handlar om brist på evidens. Det är svårt att bedöma om två behandlingar är lika, en förutsättning för klasseffekt. Bristen på bra studier som ger vägledning är också mycket stor.

Frågan handlar även om biverkningar. Det finns många exempel på biverkningar som gjort att studier avbrutits och att läkemedel dragits bort från marknaden. En besvärande underrapportering av biverkningar i kontrollerade studier tycks ske. Sådana systemfel måste snabbt repareras för att skapa balans i våra bedömningar.

Frågan handlar slutligen om den hastighet med vilken den medicinska professionen – rätt eller fel – tar sig an nya behandlingsmetoder, t ex nya trombocyttaggregationshämmare, demensläkemedel och s k COX 1-sparande NSAID-preparat. Enskild läkares förskrivning av läkemedel kan bara bli evidensbaserad om denne kontinuerligt blir uppdaterad beträffande de senaste studierna, t ex genom sjukvårdsbaserade, producentbundna aktiviteter [89]. Läkarkåren bör i högre grad efterfråga sådana.

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Furberg CD, Herrington DH, Psaty B. Are drugs within a class interchangeable? *Lancet* 1999;354:1202-4.
2. McAlister FA, Laupacis A, Wells GA, Sackett DL. User's guides to the medical literature. XIX. Applying clinical trial results. B. Guidelines for determining whether a drug is exerting (more than) a class effect. *JAMA* 1999;282:1371-7.
3. Furberg CD, Pitt B. Are all angiotensin-converting enzyme inhibitors interchangeable? *J Am Coll Cardiol* 2001;37:1456-60.
4. Massel D. Similar, the same or just not different: a guide for deciding

whether treatments are clinically equivalent. *Can J Cardiol* 1999;15:556-62.

12. Mulrow C. Evidence based medicine 2001;6 Jan-Feb:10-1.
13. Bandolier – Class and equivalence. *Bandolier* 2002;95-3. Tillgänglig på: <http://www.jr2.ox.ac.uk/bandolier/band95/b95-3.html>
18. Wright JM. Choosing a first line drug in the management of elevated blood pressure: What is the evidence? 2: β -blockers. *CMAJ* 2000;163:188-92.
19. Wright JM. Choosing a first line drug in the management of elevated blood pressure: What is the evidence? 3: Angiotensin-converting-enzyme inhibitors. *CMAJ* 2000;163:293-6.
23. Pahor M, Psaty BM, Alderman MH, Applegate WB, Williamson JD, Cavazzini C, et al. Health outcomes associated with calcium antagonists compared with other first-line antihypertensive therapies: a meta-analysis of randomised controlled trials. *Lancet* 2000; 356: 1949-54.
24. Neal B, MacMahon S, Chapman N. Effects of ACE inhibitors, calcium antagonists, and other blood-pressure-lowering drugs: results of prospectively designed overviews of randomised trials. *Blood Pressure Lowering Treatment Trialists' Collaboration*. *Lancet* 2000;356:1955-64.
58. Ioannidis J, Lau J. Completeness of safety reporting in randomized trials. An evaluation of 7 medical areas. *JAMA* 2001;285:437-43.
59. Goldkind L, Witter J, Hong L. Statistical reviewer briefing document for advisory committee. Tillgängligt på: http://www.fda.gov/ohrms/dockets/ac/01/briefing/3677b1_04_stats.doc.
60. FDA advisory committee briefing document. http://www.fda.gov/ohrms/dockets/ac/01/briefing/3677b2_03_med.doc.
66. Rubinstein E. History of quinolones and their side effects. *Chemotherapy* 2001;47 Suppl 3:3-8.
70. van Walraven C, Mamdani MM, Wells PS, Williams JI. Inhibition of serotonin reuptake by antidepressants and upper gastrointestinal bleeding in elderly patients: retrospective cohort study. *BMJ* 2001; 323:655-8.
76. Meadows M. Serious liver injury. Leading reason for drug removals, restrictions. *FDA Consumer* 2001;35:8-9.
86. Thomas M, Mann J. Increased thrombotic vascular events after change of statin. *Lancet* 1998;352:1830-1.
87. Bourgault C, Elstein E, Le Lorier J, Suissa S. Reference-based pricing of prescription drugs: exploring the equivalence of angiotensin-converting-enzyme inhibitors. *CMAJ* 1999;161:255-60.
88. Diggory P, Fernandez C, Humphrey A, Jones V, Murphy M. Comparison of elderly people's technique in using two dry powder inhalers to deliver zanamivir: randomised controlled trial. *BMJ* 2001; 322:577-9.
89. Drummond G. Are drugs interchangeable [letter]? *Lancet* 2000; 355:316.

I Läkartidningens elektroniska arkiv
<http://tarkiv.lakartidningen.se>
 är artikeln kompletterad med fullständig referenslista.

SUMMARY

Ranking the efficacy of drugs – a question of »class«? It is difficult to determine if drugs within a class are interchangeable

Hans Liedholm, Agneta Björck Linné
Läkartidningen 2002;99:3492-3502

Criteria for equivalence are essential when classifying drugs. Results from clinical trials remain the most practical way of defining »class«. Class effect cannot be assumed, it must be proven, and one should not attempt to generalize results from one drug to another drug in the same class.

Correspondence: Hans Liedholm, Dept of Clinical Pharmacology, Universitetssjukhuset MAS, SE-205 02 Malmö, Sweden (hans.liedholm@smi.mas.lu.se)



**S E U P P
M E D
S T A T I S T I K E N**

ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Adam Taube, professor, statistiker, institutionen för informationsvetenskap, Uppsala universitet, (adam.taube@dis.uu.se)

Jörgen Malmquist, docent i internmedicin, Höllviken, (jorgen.malmquist@mailbox.swipnet.se)

Räkna med vad du tror

Bayes' sats i diagnostiken

II Grundtanken i evidensbaserad medicin är att tillgängliga vetenskapliga resultat skall komma till direkt användning i det praktiska kliniska arbetet. En ansats som då är relevant är s k bayesiansk inferens som ger en struktur för att kombinera ny kunskap med det man redan vet, eller tror sig veta. Det kan hävdas att läkare i alla tider i sin kliniska verksamhet intuitivt arbetat och tänkt som »bayesianer« utan att egentligen vara medvetna om detta.

Anta att en kvinna söker akut för värk i ena underbenet. Värken började för ungefär ett dygn sedan. Klinisk undersökning ger inga starka hållpunkter för djup ventrombos, och nivån av D-dimer i plasma ligger nära övre referensområdesgränsen. Läkaren bedömer sannolikheten för djup ventrombos vara liten och överväger att uppmana patienten att komma tillbaka nästa dag för ny undersökning. Först ställer han emellertid ytterligare några frågor och får då veta att patienten använder p-piller och är rökare. Läkaren vet därmed att sannolikheten för djup ventrombos är betydligt större än vad han först tänkte sig. Han bestämmer sig därför för att tillråda patienten att flebografi utförs. Detta illustrerar hur läkaren dels förknippar en viss diagnos med en sannolikhet (utan att nödvändigtvis ha ett bestämt numeriskt värde på denna), dels utnyttjar ytterligare information för att modifiera denna sannolikhet.

I den klassiska (frekventistiska) statistiken förknippas sannolikhetsbegreppet med en situation som kan upprepas – man gör slumpmässiga dragningar av individer, observationer, ur en viss population och uttalar sig sedan om den bakomliggande populationen på grundval av dessa observationer. I den bayesianska statistiken används sannolikheter även för att mäta graden av tilltro, och man integrerar därvid information från nya observationer med tidigare tillgänglig kunskap eller gissningar. Fundamental i denna metodik är Bayes' sats som vi skall försöka presentera här.

Prästen Thomas Bayes verkade i 1700-talets England. Han var inte bara präst utan också matematiker och Fellow of the Royal Society. Efter hans död 1761 fann man hans nu berömda text om sannolikheter, »Essay towards solving a problem in the doctrine of chances«, som publicerades av Royal Society 1763. I texten fanns bland annat det som kommit att kallas Bayes' sats (eller Bayes' teorem). Bayes' tankar blev kända och omstridda, föll sedan nästan helt i glömska för att återupptäckas för några decennier sedan. Inte minst bland brittiska statistiker debatteras för- och nackdelar med bayesi-

SAMMANFATTAT

Bayesiansk analys används i växande omfattning inom medicinsk statistik, diagnostiskt tänkande och beslutsanalys.

Bayes-analysen ger en matematisk metod för integrerad bedömning av förhandskunskap och ny information. Detta tillvägagångssätt har i många sammanhang överlägsen förmåga att ge tillförlitlig uppdaterad kunskap och optimalt beslutsunderlag.

Artikeln, som är den första av två, presenterar den grundläggande sats som är utgångspunkten när bayesianskt tänkande används i diagnostiskt arbete.

Evidensbaserad medicin



Thomas Bayes (1702 – 1761). Hans teorem om betingad sannolikhet blev 200 år efter hans död utgångspunkten för betydelsefullt nytänkande inom statistisk analys.

Tabell I. Resultat och beslut vid cytologisk screening för cervixcancer.

Undersökningens utfall	Det sanna förhållandet: den undersökta kvinnan	
	har cervixcancer	har inte cervixcancer
Positivt: cellförändringar påvisade	<i>Sant positivt resultat</i> som leder till korrekt beslut: ytterligare undersökningar och eventuell behandling	<i>Falskt positivt resultat</i> som leder till fel beslut: onödiga undersökningar utförs
Negativt: inga cellförändringar	<i>Falskt negativt resultat</i> som leder till fel beslut: behövlig ytterligare undersökning blir inte utförd	<i>Sant negativt resultat</i> som leder till korrekt beslut: ytterligare (onödig) undersökning utförs inte

anskt statistiskt tänkande, bland annat inom medicinsk statistik och klinisk beslutsteori.

Den enklaste typen av sannolikhet i detta sammanhang kan illustreras med att en person väljs helt slumpmässigt ur en population. Sannolikheten att denna person skall ha en viss sjukdom är då helt enkelt densamma som prevalensen för (förekomsten av) denna sjukdom i den aktuella populationen. För det mesta rör det sig i det kliniska arbetet mer om s k betingade sannolikheter, dvs sannolikheter som påverkas av att speciella omständigheter, betingelser, föreligger. Det kan t ex vara så att det finns en viss basal (liten) sannolikhet för att en person med halsont och svullna halskörtlar har mononukleos, men om blodbilden ser ut på ett visst sätt är sannolikheten plötsligt mycket större för att den sjukdomen skall föreligga. Betingelsen här var just blodbilden.

Bayes' sats ger en metod för att modifiera den första grundläggande sannolikheten för t ex att en viss sjukdom föreligger när nya betingelser kommer in i bilden. Betingelsen kan vara utfallet av en diagnostisk åtgärd, information från patienten eller kunskap om nya forskningsresultat. Uttryckt i tekniska termer innebär detta att en a priori-sannolikhet kan omräknas till en a posteriori-sannolikhet. Denna distinktion är inte aktuell i klassisk inferensmetodik.

I engelska texter används termerna prior probability och posterior probability. Här kommer vi att använda »sannolikhet före« och »sannolikhet efter«. Motsvarande termer används när sannolikhet byts ut mot odds.

En numerisk, fiktiv illustration av Bayes' sats

Anta att man vill utföra en allmän hälsoundersökning, screening, för att i en viss population av kvinnor på ett tidigt stadium upptäcka så många fall av cervixcancer och förstadium (samtliga benämns här cervixcancer) som möjligt. Det cytologiska screeningprov som skall användas resulterar för varje undersökt person i ett provsvar, kallat »positivt« (+) när cellförändringar konstaterats och »negativt« (–) när inga sådana avvikande celler påvisats. Dessvärre fungerar inte screeningproceduren med hundraprocentig säkerhet – fel i bägge riktningarna kan förekomma. Se Tabell I.

För att kvalitetsbestämma en diagnostisk procedur brukar man ange två betingade sannolikheter, där betingelsen bestäms av huruvida personen har den aktuella sjukdomen eller ej.

Sensitivitet: $P(+|S)$ = sannolikheten för positivt testutfall hos sjuk person

Specificitet: $P(-|ej S)$ = sannolikheten för negativt testutfall hos ej sjuk person

I notationerna står P för probabilitet, dvs sannolikhet, och (+|S) för »positivt testutfall givet att personen har sjukdomen«. Sannolikheter uttrycks som tal mellan 0 och 1.

Storheterna sensitivitet och specificitet lanserades av Yerushalmy i ett klassiskt arbete 1947 [1]. Om dessa sannolikheter är stora visar detta att den aktuella proceduren är effektiv. Idealfallet är givetvis att båda två har värdet 1, dvs att summan av sensitivitet och specificitet är 2. Om summan är 1 är den diagnostiska proceduren helt värdelös.

Sensitivitet och specificitet kan skattas på stora serier av individer med respektive utan den aktuella sjukdomen. Beträffande vårt exempel kan representativa siffror anses vara: sensitivitet 0,85 och specificitet 0,95. Kännedom om dessa siffror är av värde för berörda läkare, men det räcker inte. Det som är av störst intresse är rimligen dels hur många av kvinnorna med positivt provsvar (som kräver ytterligare undersökning) som verkligen har sjukdomen, dels hur många av dem som »frikänns« på grund av negativt provsvar som verkligen inte har sjukdomen. Detta är en annan typ av betingade sannolikheter, där betingelsen bestäms av provsvaren, nämligen:

Positivt prediktionsvärde:

$P(S|+)$ = sannolikheten för sjukdom hos person med positivt testutfall

Negativt prediktionsvärde:

$P(ej S|-)$ = sannolikheten för »ej sjukdom« hos person med negativt testutfall

Dessa två storheter kan inte beräknas på grundval av skilda serier av personer med respektive utan den aktuella sjukdomen. De är i stället knutna till en speciell population där sjukdomen har en viss förekomst (prevalens), $P(S)$. Anta att den aktuella populationen består av 10 000 kvinnor och att prevalensen av cervixcancer är 0,01. Prevalenstalet är i bayesiansk terminologi »före-sannolikheten«. I populationen finns alltså 100 kvinnor med cervixcancer. Bland dessa förväntas $0,85 \times 100 = 85$ stycken positiva provsvar (sant positiva). Dessvärre kan det dessutom förväntas ett antal positiva provsvar bland dem som inte har sjukdomen (falskt positiva), i detta fall $0,05 \times 9\,900 = 495$ stycken.

Det positiva prediktionsvärdet blir alltså $85/(85+495) = 0,147$ eller cirka 15 procent. Denna siffra kan förefalla förvånande låg. Förklaringen är givetvis att prevalensen för sjukdomen är så låg. Hade prevalensen varit högre hade det blivit fler sant positiva och färre falskt positiva utfall. Motsvarande kalkyl för det negativa prediktionsvärdet blir följande. Sant negativa $0,95 \times 9\,900 = 9\,405$, falskt negativa 15, negativt prediktionsvärde $9\,405/9\,420 = 0,998$. Sannolikheten för att en kvinna som får negativt testutfall verkligen inte har cervixcancer är alltså mycket nära 1.

Prediktionsvärdenas beroende av prevalensen vid sensitivitet 0,85 och specificitet 0,95 illustreras av Figur 1.

Prediktionsvärdena utgör anknytningen till Bayes' sats, i den utformning den har givits för att användas i det medi-

II Fakta 1

Bayes sats tillämpad på diagnostik

$P(S)$: sannolikheten före, dvs sannolikheten för att en viss person skall ha sjukdomen. Denna sannolikhet svarar mot sjukdomens prevalens.

$P(+|S)$: sannolikheten för att en person med sjukdomen får ett positivt testutfall (testets *sensitivitet*).

$P(-|ej S)$: sannolikheten för att en person som inte har sjukdomen får ett negativt testutfall (testets *specificitet*).

$P(S|+)$: sannolikheten efter, här sannolikheten för sjukdom vid positivt testutfall (testets **positiva prediktionsvärde**).

$P(ej S|-)$: sannolikheten efter, här sannolikheten för icke-sjukdom vid negativt testutfall (testets **negativa prediktionsvärde**).

$$P(S|+) = \frac{P(+|S) \times P(S)}{P(+|S) \times P(S) + [1 - P(-|ej S)] \times [1 - P(S)]}$$

$$P(ej S|-) = \frac{P(-|ej S) \times [1 - P(S)]}{P(-|ej S) \times [1 - P(S)] + P(-|S) \times P(S)}$$

Bayes sats med användning av odds

Sannolikheterna för resp emot sjukdom omvandlas till odds:

$$\text{Odds före: } \frac{P(S)}{P(ej S)} \quad \text{Odds efter: } \frac{P(S|+)}{P(ej S|+)}$$

Med hjälp av Bayes-ekvationen får vi följande samband vid positivt testutfall:

$$\frac{P(S|+)}{P(ej S|+)} = \frac{P(S)}{P(ej S)} \times \frac{P(+|S)}{P(+|ej S)}$$

där det avslutande bråket är **likelihood-kvoten** vid positivt testutfall, **LR+**.

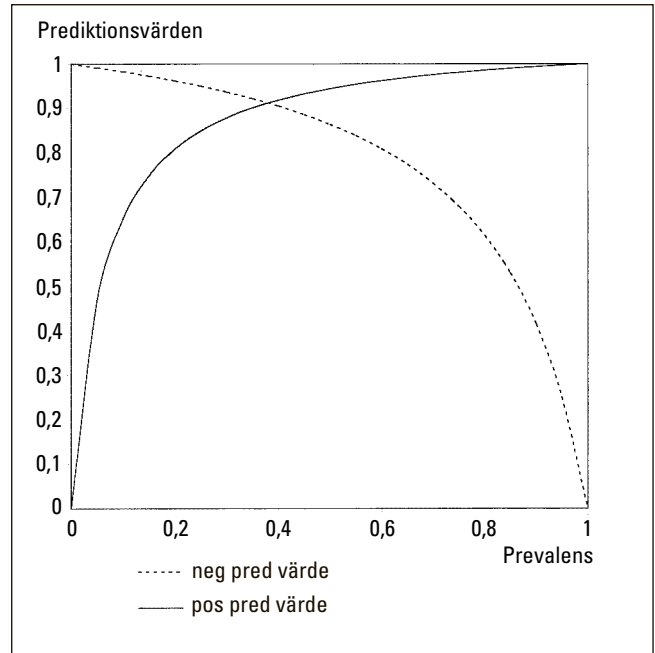
Alltså: $\text{odds efter} = (\text{odds före}) \times (\text{likelihood-kvoten})$.

Märk att **LR+** = (sant positiva utfall) / (falskt positiva utfall).

cinskt diagnostiska sammanhanget. Den ger möjlighet att utifrån prevalensen (före-sannolikheten för sjukdom) samt de betingade sannolikheterna (sensitivitet och specificitet) som beskriver testets effektivitet beräkna de betingade sannolikheterna (efter-sannolikheterna) där provsvaren utgör betingelserna. Prediktionsvärdena är alltså identiska med efter-sannolikheten för sjukdom vid positivt testutfall respektive efter-sannolikheten för icke-sjukdom vid negativt testutfall. Proceduren återges formaliserad i Fakta 1.

Att räkna med odds

Den struktur vi just presenterat i termer av betingade sannolikheter kan även uttryckas med hjälp av odds. Om sannolikheten för en händelse (H) betecknas med $P(H)$ är oddset för denna händelse kvoten $P(H) / P(ej H) = P(H) / [1 - P(H)]$. Man kan t ex säga att oddset för att få ett hjärterkort vid dragning av ett kort ur en kortlek är »ett mot tre«.



Figur 1. Inverkan av prevalensen på det positiva prediktionsvärdet och det negativa prediktionsvärdet då den diagnostiska metodens sensitivitet är 0,85 och specificitet 0,95.

I det fiktiva exemplet ovan är oddset för att en ur hela populationen slumpvis utvald kvinna har cervixcancer »ett mot nittionio« (före-odds). Om det gäller en kvinna som fått ett positivt provsvar blir efter-oddset kvoten mellan de betingade sannolikheterna uträknade enligt Bayes' sats, dvs

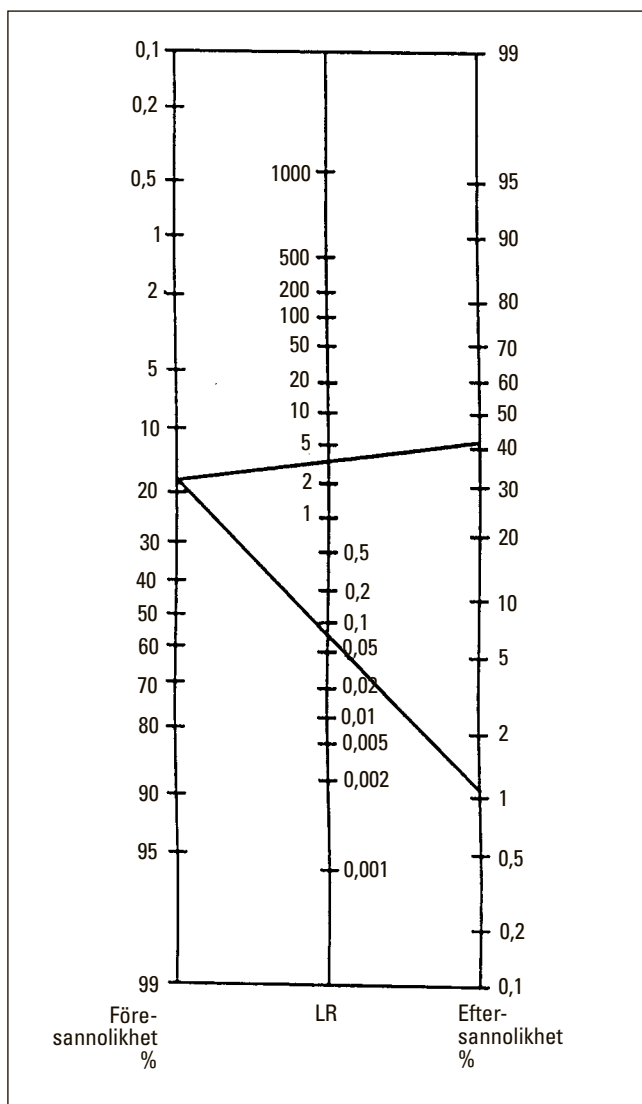
$$\frac{P(S|+)}{P(ej S|+)} = \frac{85/(85 + 495)}{495/(85 + 495)} = \frac{17}{99}$$

Efter-oddset blev alltså 17 gånger högre än före-oddset. Denna multiplikationsfaktor svarar mot kvoten sensitivitet / (1 - specificitet) = 0,85 / (1 - 0,95). Denna kvot brukar på klingande svenska kallas likelihood-kvoten och förkortas LR (för likelihood ratio). (Det engelska ordet likelihood betyder sannolikhet eller trolighet, men brukar användas oöversatt för att den speciella innebörden skall framgå.)

Alltså: om man multiplicerar före-odds för sjukdom med likelihoodkvoten får man efter-odds för sjukdom. Se också Fakta 1. Eftersom likelihoodkvoten lätt kan kalkyleras ur testmetodens prestanda erbjuder oddsmetoden en tilltalande enkel metod att beräkna efter-sannolikheten för sjukdom. Det erhållna efter-oddset kan nämligen lätt omvandlas till efter-sannolikhet med hjälp av sambandet: sannolikhet = odds / (1 + odds). I det här fallet är efter-odds 17/99 = cirka 0,17 och efter-sannolikheten 0,147 eller cirka 15 procent, samma värde som vi erhöll tidigare.

Resonemanget gäller i detta fall likelihoodkvoten för positivt testutfall, **LR+**. Man kan på motsvarande sätt beräkna efter-odds för att sjukdom inte föreligger vid negativt testutfall och använder då **LR-** som är (1 - sensitivitet) / specificitet. Användningen av odds och LR behandlas i en bok om diagnosstrategier [2].

LR kan användas som en mätare på värdet av ett diagnostiskt kriterium. **LR+** anger hur många gånger vanligare ett positivt utfall är bland dem som har sjukdomen än bland dem som inte har sjukdomen. Diagnoskriteriet är alltså av större värde ju högre **LR+** -siffran är. Omvänt är **LR-** mera vägledande ju närmare noll det ligger. En illustration ges av den undersökning av olika diagnoskriteriers **LR**-värden för dia-



Figur 2. Nomogram för omvandling av före-sannolikhet till efter-sannolikhet. Exemplet gäller analys av D-dimer i diagnostiken av ventrombos. Analysens likelihood-kvoter är vid positivt resultat (LR+) 3,24 och vid negativt resultat (LR-) 0,08. Sannolikheterna är i detta fall uttryckta i procenttal. Före-sannolikheten (prevalensen) är 17 procent. Man kan avläsa att ett positivt analysresultat ger efter-sannolikhet drygt 40 procent, medan ett negativt resultat ger efter-sannolikhet ca 1 procent. (Från Eliasson [4] med författarens medgivande.)

gnostik av appendicit som utförts av Andersson och medarbetare [3].

Flertalet diagnostiska metoder har känd sensitivitet och specificitet, som har för metoden konstanta värden (vi förbigår här att detta inte alltid gäller oinskränkt). Alltså har varje sådan metod kända värden för LR+ och LR-, och det är väl motiverat att använda dessa enligt ovan för att ge det diagnostiska arbetet större precision. Ett nomogram kan vara till hjälp när man med LR-värden vill omvandla före-sannolikheter till efter-sannolikheter (Figur 2).

Delade meningar

Bayes' sats så som den presenterats i vårt fiktiva exempel ovan inbjuder knappast till några större tolkningskontroverser. Man behöver inte vara »bayesian« för att acceptera vad som sägs där under förutsättning att det rör sig om just en population där förekomsten av cervixcancer verkligen är en på hundra.

Om man inte har kunskap om prevalensen utan endast känner till testprocedurens sensitivitet och specificitet kan man naturligtvis ändå *tro* eller *gissa* något om den verkliga förekomsten av cervixcancer i populationen. Anta att det gäller en kvinna som fått ett positivt provsvar (cellförändringar) och att man vill ange en sannolikhet för att hon har cervixcancer. Om man t ex tror att prevalensen av cervixcancer är 1/200 och sätter in detta i uttrycket för $P(S|+)$ så blir sannolikheten för att en kvinna med positivt provsvar har cervixcancer knappt 8 procent i stället för nära 15 procent som ovan. Givetvis får man samma resultat om man använder oddsmetoden, dvs multiplicerar före-odds (0,005/0,995) med metodens LR+ - värde 17 för att få efter-odds som sedan konverteras till efter-sannolikhet.

Det som gjort tillämpningen av Bayes' metodik så omdiskuterad är att man vid beräkningen gällande ett enskilt fall utnyttjar inte enbart iakttagelser om denna individ (»kvinna fick ett positivt testsvar«) utan även inkluderar gissade storheter (»förekomsten av cervixcancer är nog 1/200«) vid beräkningen av ett mått för tilltron till den ställda diagnosen.

Vi har här försökt illustrera den bayesianska ansatsen med ett mycket enkelt exempel. Om man accepterar fundamentet för denna metod kan den emellertid utvecklas i många olika riktningar och bli synnerligen användbar i ett flertal medicinska sammanhang. I en följande artikel kommer vi att presentera sådana ytterligare tillämpningar av Bayes-analys.

Referenser

1. Yerushalmy J. Statistical problems in assessing methods of medical diagnosis, with special reference to X-ray techniques. *Pub Health Rep* 1947;62:1432-49.
2. Suchman AL, Dolan JG. Odds and likelihood ratios. In: Black ER, Bordley DR, Tape TG, Panzer RJ, editors. *Diagnostic strategies for common medical problems*. 2nd ed. Philadelphia: American College of Physicians; 1999.
3. Andersson R, Hugander A, Ghazi S, Ravn H, Offenbartl K, Nyström PO, et al. Diagnostic value of disease history, clinical presentation, and inflammatory parameters of appendicitis. *World J Surg* 1999; 23:133-40.
4. Eliasson M. Att använda diagnostiska test. I: Asplund K, Berglund G, Lindgren G, Lindholm N, red. *Internmedicin 2000* (manus).

SUMMARY

Count on your beliefs Bayes' theorem in diagnosis

Adam Taube, Jörgen Malmquist
Läkartidningen 2001; 98: 2910-3

Bayesian analysis of data finds increasing use in medical statistics, diagnostic evaluation and decision analysis. The central element in bayesian analysis is a set of mathematical rules for integrated evaluation of prior knowledge and new information. In many situations this approach has superior ability to deliver dependable updated knowledge and to provide an optimal probability basis for decisions. This article (the first of two) presents Bayes' theorem and its application in diagnostic work. It is explained how likelihood ratios of diagnostic tests interact with the outcome of such tests in the conversion of initial information (prior odds) to enhanced information (posterior odds).

Correspondence: Adam Taube, Dept of Information Science, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden. (adam.taube@dis.uu.se)

Adam Taube, professor, statistiker, institutionen för informationsvetenskap, Uppsala universitet (*adam.taube@dis.uu.se*)

Jörgen Malmquist, docent i internmedicin, Höllviken (*jorgen.malmquist@mailbox.swipnet.se*)

Räkna med vad du tror

Bayes – inte P-värdet – mäter tilltron

II Det är tyvärr alldeles för vanligt att medicinska forskare inte är medvetna om att varje statistisk analys grundas på en teoretisk modell, och att de slutsatser som kan dras från en statistisk analys av data gäller under just denna modell. Valet av analysmodell styr därför alltid resultatet – anpassar man t ex en rät linje till tvådimensionella data blir resultatet just en rät linje och inget annat, oavsett om detta är en realistisk beskrivning eller ej. När det gäller situationer där signifikansanalys och konfidensintervall brukar användas kan valet av en bayesiansk modell ibland leda till fundamentalt annorlunda tolkningar än den traditionella sk frekventistiska analys som för närvarande helt dominerar den medicinska litteraturen.

Vid en skattning av t ex risken för ventrombos hos kvinnor som använder p-piller förutsätter den traditionella, frekventistiska modellen att en »sann« risk (med ett exakt numeriskt värde, traditionellt betecknat med en grekisk bokstav) existerar och att de data man insamlar i en viss undersökning ger ett estimat av denna bakomliggande fixa okända parameter. Den frekventistiska ansatsen, som förklaras i följande stycke, har under decennier välsignat medicinska artiklar med ett ymnigt regn av signifikansstjärnor och P-värden som inte sällan fel-tolkats och ibland till och med saknat relevans [1].

Tolkningen av ett konfidensintervall

Anta t ex att man beräknat ett 95 procents konfidensintervall för patienternas genomsnittliga blodtryckssänkning vid en klinisk prövning av ett nytt blodtryckssänkande läkemedel, och att detta intervall blev $6,2 \pm 4,7$ mm Hg. Den korrekta, traditionella tolkningen är då: »detta intervall hade (märk imperfektum) sannolikheten 95 procent att falla så att det täckte det sanna (men okända) värdet för blodtryckssänkningen«. Den modell som hela kalkylen baseras på förutsätter alltså att det existerar ett enda sant, fixt värde som man söker estimera så bra som möjligt. Resonemanget utgår från vad som kan förväntas i det långa loppet: om man skulle upprepa undersökningen många gånger och varje gång beräkna ett 95 procents konfidensintervall från de erhållna data skulle i genomsnitt 19 av 20 sådana intervall falla så att de innefattar »det sanna värdet« (därav termen »frekventistisk«). Man har därmed 95 procents konfidens (förtroende) för detta intervall.

Inte sällan förekommer i medicinska sammanhang en annan, felaktig formulering som i det här fallet skulle lyda: »med 95 procents säkerhet är det sant att den verkliga genomsnittliga blodtryckssänkningen ligger i intervallet $6,2 \pm 4,7$ «. Eftersom det aktuella intervallet begränsas av fixa numeriska

SAMMANFATTAT

Bayesiansk analys av medicinska data beskrivs och exemplifieras i denna artikel, som är den andra av två.

Grundtanken i traditionell (frekventistisk) statistisk analys förklaras med utgångspunkt i begreppet konfidensintervall.

Bayesiansk analys i diagnostik exemplifieras: i valet mellan olika diagnoser sammanvägs tillståndens relativa förekomst i patientpopulationen med sensitiviteterna för ett diagnostiskt test.

Skillnaderna mellan traditionell och bayesiansk analys av resultaten av en klinisk prövning beskrivs, och fördelarna hos den senare förklaras.

En kort litteratursammanfattning visar den utbredda tillämpbarheten av bayesiansk analys inom medicinsk statistik.

Evidensbaserad medicin

En tidigare artikel om Bayes-analys har publicerats i Läkartidningen 24/2001.



Thomas Bayes (1702–1761).

värden förutsätter denna formulering att det sanna värdet på något sätt skulle vara variabelt. Detta leder emellertid till en helt annan teoretisk modell och avspeglar väl närmast hur många medicinare skulle önska att ett dylikt intervall kunde

Tabell I. Bayesiansk kalkyl vid diagnostik med hjälp av cancermarkören CA 125.

1 Sjukdomskategori	2 Andel bland ifrågavarande patienter (före-sannolikhet)	3 Sensitivitet hos CA 125-analys (likelihood)	4 Produkten av 2 och 3	5 Efter-sannolikhet
Ovarialcancer	0,40	0,85	0,34	0,73
Annan cancer	0,10	0,50	0,05	0,11
Benign tumör	0,50	0,15	0,075	0,16
Summa	1,00		0,465	1,00

tolkas. Denna tolkning stämmer, som skall beröras ytterligare, snarare överens med den bayesianska modellen än med den frekventistiska.

Sammanvägning av tillgänglig information

Vår föregående artikel [2] presenterade Bayes' sats i en situation med dikotomier – »sjuk/ej sjuk« och »positivt/negativt testutfall«. Här skall strukturen belysas med ett fiktivt exempel där flera sjukdomar är aktuella. Vi utgår från kvinnor som vid gynekologisk undersökning befunnits ha en tumörmiss-tänkt förändring intill uterus. Undersökningen kompletteras med analys av tumörantigenet CA 125 i serum. Förhöjd nivå av denna indikator antas ha följande sensitivitetvärden: vid ovarialcancer omkring 85 procent och vid andra tänkbara cancerformer ca 50 procent. Vidare antas 15 procent av patienter med benign tumör ha förhöjd CA 125-nivå. Anta vidare att man vid kliniken har erfarenheten att bland patienter med palpationsfyndet ifråga 40 procent har ovarialcancer, 10 procent annan cancer och 50 procent benign tumör. Detta innebär bl a att man i avsaknad av ytterligare diagnostisk information anser sannolikheten vara 40 procent för att en patient med detta palpationsfynd har ovarialcancer.

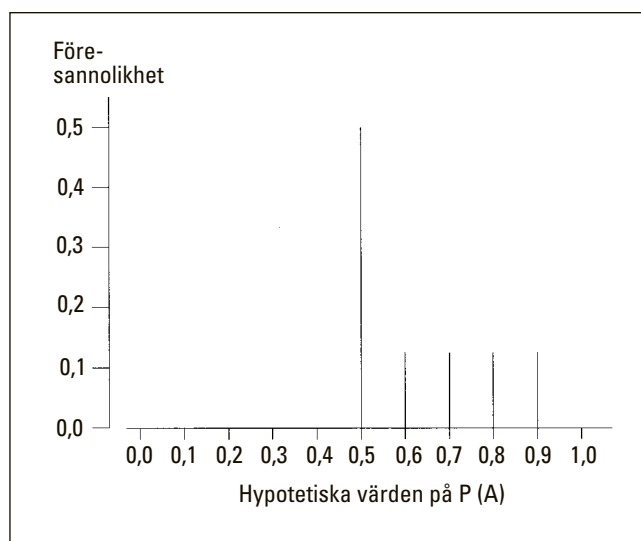
Bland dessa patienter blir andelen kvinnor med förhöjt CA 125-värde sammanlagt $0,85 \times 0,4 + 0,5 \times 0,1 + 0,15 \times 0,5 = 0,465$ och andelen som har ovarialcancer samt förhöjt CA 125-värde $0,85 \times 0,4 = 0,34$, se Tabell I. Sannolikheten för att en viss kvinna med förhöjt CA 125 har ovarialcancer blir givetvis andelen ovarialcancerfall bland alla dem som har förhöjt CA 125, dvs $0,34/0,465 = 0,73$.

Utän att beakta annan diagnostisk information skulle man alltså på grundval av palpationsfynd och förhöjt CA 125 kunna hävda att sannolikheten är 73 procent att denna kvinna har ovarialcancer. Denna kalkyl grundas på uppfattningen om proportionerna (0,4 respektive 0,1 och 0,5) mellan de tre diagnoskategorierna. Dessa proportioner har vägts samman med respektive sensitivitet. Varje sensitivitetstal anger troligheten (likelihood) att en person inom respektive sjukdomskategori skall komma med bland dem som har ett positivt CA 125-värde.

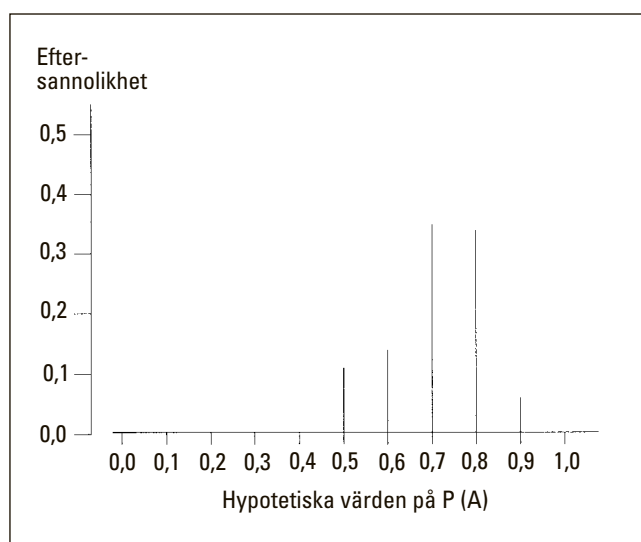
I detta exempel ändrades alltså före-sannolikheten (prior probability) 40 procent genom det förhöjda CA 125-värdet till efter-sannolikheten (posterior probability) 73 procent. Det är uppenbart att denna kalkyl är starkt beroende av de nämnda erfarenhetsmässiga proportionerna mellan olika patientkategorier; skulle man ändra dessa före-sannolikheter skulle slutresultatet bli ett annat. Det kan också vara fallet att sådana värden inte alls är kända och att berörda läkare endast har på känn ungefär hur stora de kan vara. Men även detta är kunskap.

Traditionell »frekventistisk« analys

Anta att man vill jämföra två migränpreparat genom en överkorsningsstudie där 20 migränpatienter får ange vilket preparat de föredrar utifrån bästa symtomlindrande effekt. Av de båda medlen A och B innehåller B ett standardpreparat, me-



Figur 1a. A priori-fördelning av sannolikheterna för att olika andelar av patienterna föredrar läkemedel A i en prövning av A mot B.



Figur 1b. A posteriori-fördelning av sannolikheterna för att olika andelar av patienterna föredrar läkemedel A. Dessa efter-sannolikheter är beräknade med ledning av prövningsresultatet, där 15 av 20 patienter föredrog A. Se Tabell II.

dan A innehåller samma preparat plus en tillsats i form av en ny substans. Den senares eventuella förmåga att förstärka effekten skall studeras. Det anses finnas goda skäl att anta att tillsatsen inte orsakar försämrade effekt, men det är ovisst om den har någon positiv verkan. Nollhypotesen enligt gängse analysmodell är att medlen är lika effektiva, vilket innebär att det kan förväntas att antalet patienter som föredrar A är detsamma som antalet som föredrar B. Detta innebär att sanno-

Tabell II. Bayesiansk resultat kalkyl vid överkorsningsprövning av läkemedlen A och B, varvid 15 av 20 patienter föredrog A.

1 Hypotes (andel av patienter som föredrar A)	2 Före-sannolikhet	3 Likelihood ($\times 100\ 000$)	4 Produkten av 2 och 3 ($\times 100\ 000$)	5 Efter-sannolikhet
0,5	0,500	0,0954	0,0477	0,114
0,6	0,125	0,4815	0,0602	0,144
0,7	0,125	1,1537	0,1442	0,345
0,8	0,125	1,1258	0,1407	0,336
0,9	0,125	0,2059	0,0257	0,061
Summa	1,000		0,4185	1,000

likheten för att föredra A är $P(A) = 0,5$. Om A har större effekt än B skall man förvänta sig att $P(A) > 0,5$. Man beslutar att resultatet skall signifikansprövas på 5-procentsnivån med enkelsidigt test.

Resultatet blev att 15 patienter föredrog A och fem föredrog B. Detta ger ett s k P-värde som är 0,02 (enligt exakt binomialfördelning med $n=20$ och $\pi=0,5$). Det man nu vet är alltså: »sannolikheten för att få det resultat som verkligen erhöles (eller något ännu extremare) var så liten som 0,02 – om nollhypotesen verkligen varit sann«. Observera att detta inte innebär att sannolikheten för att nollhypotesen är sann är 0,02.

Konfidensintervallet (95 procent) för proportionen patienter som föredrar A blir 0,51 till 0,91. Ett 95 procents konfidensintervall, som ju innehåller alla tänkbara hypotetiska värden på $P(A)$ som inte kan förkastas på grundval av aktuella data, täcker alltså här nästan hela skalan av värden större än 0,5. Detta innebär att man har 95 procents konfidens för det aktuella intervallet men ingen bestämd sannolikhet för att enskilda värden är sanna, även om den funna proportionen $15/20 = 0,75$ är en mycket rimlig gissning som den mest troliga (maximum likelihood-estimat).

Analys enligt Bayes

Vi utgår från samma prövning som i föregående avsnitt. Anta att prövarna först tillfrågade en rad kolleger om vad de trodde om möjligheten att den nya tillsatsen skulle ha en effektförbättrande verkan. Kollegerna fick helt enkelt, på grundval av sin kunskap, erfarenhet och intuition, föreslå (gissa) ett numeriskt värde på $P(A)$, för enkelhets skull angivet endast som hela tiondelar, dvs 0,1, 0,2 etc. Det visade sig att ungefär hälften inte trodde att den nya tillsatssubstansen skulle ha någon märkbar effekt, dvs de angav som troligt värde $P(A) = 0,5$ (nollhypotesen). De som inte trodde på nollhypotesen var alla inne på att tillsatsen kunde förväntas ha en positiv effekt, dvs att värdet på $P(A)$ borde vara $>0,5$. Vidare precisering hade man svårt att enas om, men ingen trodde att tillsatsen skulle vara så effektiv att alla patienter skulle föredra medlet A. Man enades om att som alternativ till $P(A) = 0,5$ studera värdena 0,6, 0,7, 0,8 och 0,9 vilka av de berörda ansågs vara ungefär lika troliga. Den tilltro som en samlad expertis alltså tillmätte de olika alternativa värdena på $P(A)$ kunde därmed sammanfattas i den »a priori-fördelning« som illustrerats i Figur 1a.

Hur förändras nu tilltron till tillsatssubstansens effekt genom resultatet från prövningen? Betrakta t ex alternativet $P(A) = 0,6$. Då är likelihood-värdet för att man skall få just den kombination av resultat som prövningen gav

$$(0,6)^{15} \times (0,4)^5 = 0,4815 \times 10^{-5}$$

Likelihood-värdet anger alltså sannolikheten för att få det resultat som faktiskt erhöles, om just denna hypotes vore riktig (jfr sensitivitetens roll i exemplet ovan med diagnostisk

användning av CA 125!). För varje värde på $P(A)$ i a priori-fördelningen beräknas på detta sätt en likelihood, se Tabell II. Sedan sammanvägs de olika a priori-sannolikheterna med respektive likelihood. Efter justering så att totalsumman blir 1 erhålles den a posteriori-fördelning som anges i sista kolumnen i Tabell II och som illustreras i Figur 1b. Denna fördelning skildrar alltså efter-sannolikheterna, dvs den tilltro till var och en av de uppställda hypoteserna som man får fram genom en kombination av vad man trodde före prövningen (före-sannolikheterna) och det man fick fram vid själva prövningen.

A posteriori-fördelningen visar dels att det blir störst tilltro till värdena $P(A) = 0,7$ och $P(A) = 0,8$ med över 30 procent för vardera, dels att tilltron till nollhypotesen har ett värde på ungefär 11 procent. Detta skiljer sig påtagligt från det P-värde som erhöles vid den traditionella frekventistiska analysen och som mätte något annat, nämligen risken att felaktigt förkasta nollhypotesen. Skillnaden beror givetvis på att man i den bayesianska analysen har vägt in det faktum att hälften av experterna från början inte trodde att tillsatssubstansen skulle ha någon märkbar effekt.

Bayes-analys i litteraturen

Här följer några korta referat av publikationer om användning av bayesiansk analys inom medicinen.

Analys av resultat från kliniska prövningar. Ett exempel på bayesiansk värdering av prövningsdata är en ofta citerad artikel av Brophy och Joseph [3]. Den granskade rapporten från GUSTO-studien som jämförde alteplas (t-PA) med streptokinas vid hjärtinfarkt. Rapporten hade konkluderat att alteplas var överlägset och att nollhypotesen (ingen skillnad) kunde förkastas med $P=0,006$. Brophy och Joseph visar i siffror och grafik de efter-fördelningar som blev resultatet då GUSTO-data analyserades med användning av resultaten av två tidigare jämförande prövningar som före-fördelningar, vilka gavs tilltro med 0, 10, 50 eller 100 procent. De fann att det rådde betydande osäkerhet om existensen av en kliniskt relevant effektskillnad mellan preparaten.

Planering av kliniska prövningar. Vid planering av behandlingsprövningar i fas I–II är en bayesiansk ansats av värde. Man kan integrera den mycket preliminära kunskapen om ett nytt medels effekter med olika antaganden om storleken av önskvärda effekter och biverkningar, och därigenom få hållpunkter för sannolikt bästa utformning av studien, regler för monitorering och hur många personer som behöver studeras [4].

Epidemiologiska data. En artikel av Lilford och Braunholtz [5] tar sin utgångspunkt i den omdiskuterade eventuella skillnaden i risk för trombos mellan andra och tredje »generationen« av p-piller. Författarna beskriver hur Bayes-analys ger en

bättre grund än konventionell statistisk analys när man, med data från epidemiologiska undersökningar, skall bedöma sannolikheten för att en väsentlig riskdifferens finns. Artikeln ger också en introduktion till bayesianska grundbegrepp. Författarna betecknar en ökad användning av bayesianska metoder som en nödvändighet för välgrundade beslut inom »public policy«.

Medicinsk teknologivärdering. En monografi från en brittisk utvärderingsmyndighet [6] ger en grundlig beskrivning och exemplifiering av bayesiansk analys, i jämförelse med klassisk statistisk analys, inom detta område. Publikationen innehåller en beskrivande förteckning över analysprogram som kan hämtas från Internet.

Litteratur med metodologiska kommentarer. Bayes-analys berörs kort i flera böcker om medicinsk statistik, t ex det välkända verket av Armitage och Berry [7], som också hänvisar till litteratur som är helt inriktad på Bayes-metodik.

En redaktionell artikel [8] berör skillnaden mellan klassisk (frekventistisk) och bayesiansk analys, och framför skäl för ökad användning av Bayes-metodik vid bedömning av resultatet av kliniska prövningar.

Ett antal skribenter kombinerar en beskrivning av Bayes-analysens fördelar med en redogörelse för nackdelarna med fixering vid P-värden och åtföljande rigida uppdelning av forskningsresultat i »signifikanta« och »icke-signifikanta«. Det framhålls att detta binära tänkande bör ersättas av att alla data så långt möjligt tolkas inom ramen för övrigt tillgängligt vetande. Sterne och Davey Smith [9] har nyligen publicerat en artikel där schablonmässig signifikantestning kontrasteras mot bayesiansk bedömning.

En mycket grundlig och intressant genomgång av tankegångarna bakom signifikantestning, hypotesprövning och bayesiansk analys har gjorts av Goodman [10, 11]. Han ger argument för en markant kritisk inställning till icke-bayesiansk analys. Han konstaterar bland annat att ordinära metaanalyser visar tendens till bayesiansk anda genom att göra en integrerad bedömning av alla tillgängliga data, men förklarar varför de inte kan anses ge en tillräckligt djupgående analys. Goodmans artiklar sammanfattades och kommenterades positivt av redaktören för *Annals of Internal Medicine* [12].

Argument kring Bayes-analys

Bayesiansk analys ifrågasätts ibland på den grunden att det är diskutabelt att nya data modifieras av a priori-kunskaper, eftersom de senare kan vara mycket osäkra eller begränsade. Motargumentet är att det är irrationellt att bortse från befintlig kunskap eller uppfattning när man värderar nya resultat. Graden av osäkerhet hos a priori-kunskaperna kan man ta hänsyn till i den bayesianska kalkylen. Integrationen av nya data med före-sannolikheter kan utföras ett antal gånger med varierade förutsättningar. Man kan införa korrektionsfaktorer för påvisad eller förmodad bias i de tidigare eller de nya data, och man kan studera den effekt som uppkommer om vissa data modifieras (känslighetsanalys). Vidare ger Bayes-analys den värdefulla möjligheten att beräkna sannolikheten för att skillnaden mellan två behandlingsmetoders effekter överstiger ett visst värde. Man kan också beräkna s k prediktionsintervall (på engelska även kallade credible intervals). Dessa är besläktade med konfidensintervallen i klassisk analys. Ett prediktionsintervall har emellertid fördelen att ha den något annorlunda och mer lättfattliga innebörden att det sanna populationsvärdet med den beräknade sannolikheten befinner sig inom intervallet.

Kalkylarbetet kan vara mödosamt. De exempel vi använt är enkla: det rör sig om ett begränsat antal observationer och

diskreta utfallsmöjligheter (modeller). Uträkningen av tillhörande likelihoodvärden är därför enkel. Bland annat när det gäller kontinuerliga variabler i stället för diskreta kan uträkningsarbetet bli mycket omfattande. Detta är en av orsakerna till att användningen av Bayes-analys hittills varit begränsad. De vanliga kommersiellt tillgängliga mjukvarorna för statistiska kalkyler inkluderar för närvarande inte bayesianska uträkningar, men detta kommer sannolikt att förändras inom kort. Vidare finns bayesianska kalkylprogram tillgängliga på Internet, se ovan. Ett annat hindrande faktor har utgjorts av att uträkningarna kan kräva stor datorkraft, men dagens datorer har i regel tillräckliga prestanda.

Referenser

1. Taube A. Det borde (inte) vara stjärnor... *Läkartidningen* 1985;82:2422-4.
2. Taube A, Malmquist J. Räkna med vad du tror. Bayes' sats i diagnostiken. *Läkartidningen* 2001;98:2910-3.
3. Brophy JM, Joseph L. Placing trials in context using Bayesian analysis. GUSTO revisited by Reverend Bayes. *JAMA* 1995;273:871-5.
4. Thall PF, Simon RM, Estey EH. New statistical strategy for monitoring safety and efficacy in single-arm clinical trials. *J Clin Oncol* 1996;14:296-303.
5. Lilford RJ, Braunholtz D. The statistical basis of public policy: a paradigm shift is overdue. *BMJ* 1996;313:603-7.
6. Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. *Health Technol Assess* 2000; 4: No 38. Även tillgänglig via Internet: <http://www.ncchta.org>
7. Armitage P, Berry G. *Statistical methods in medical research*. 3rd ed. Oxford: Blackwell Science, 1994.
8. Freedman L. Bayesian statistical methods. A natural way to assess clinical evidence. *BMJ* 1996;313:569-70.
9. Sterne JAC, Davey Smith G. Sifting the evidence – what's wrong with significance tests? *BMJ* 2001;322:226-31.
10. Goodman SN. Toward evidence-based medical statistics. 1 The P value fallacy. *Ann Int Med* 1999;130:995-1004.
11. Goodman SN. Toward evidence-based medical statistics. 2 The Bayes factor. *Ann Int Med* 1999;130:1005-13.
12. Davidoff F. Standing statistics right side up. *Ann Int Med* 1999;130:1019-21.

SUMMARY

Count on your beliefs
Bayes – not the p value – measures credence

Adam Taube, Jörgen Malmquist

Läkartidningen 2001; 98: 3208-11

This article (the second of two) describes traditional (frequentist) statistical analysis in the context of the confidence interval.

Bayesian analysis is described in two settings. In the choice between diagnostic alternatives, the bayesian approach offers useful integration of new information with previous knowledge.

With regard to the evaluation of clinical trial data, this article exemplifies bayesian analysis as contrasted with traditional analysis, and advantages of the former are cited.

A brief literature review exposes the wide applicability of bayesian analysis in medical statistics.

Correspondence: Adam Taube, Dept of Information Science, Uppsala University, Box 513, SE-751 20 Uppsala, Sweden. (adam.taube@dis.uu.se)

Risk och odds – hur man räknar med händelser

II Medicinsk forskning innebär att man studerar variabler. De iakttagelser man gör om en variabel kallas observationer. Observationerna beskrivs med vad vi kallar data: beteckningar eller mätetal. Granskning och statistisk bearbetning av data siktar på att göra det möjligt att dra slutsatser.

Olika typer av data

Variabler av olika beskaffenhet ger olika typer av data. Grundtyperna är kategoridata (kvalitativa data) och numeriska data (kvantitativa data). Kategoridata bygger på någon form av klassifikation, t ex man/kvinna, levande/död, har viss sjukdom/har inte sjukdomen, yrke, nationalitet. Kategoridata med endast två alternativ kan kallas binära eller dikotoma. Numeriska data härstammar från variabler som kan mätas eller räknas, t ex kroppsvikt, blodglukos, antal barn i en familj.

Under bearbetningen av resultat blir insamlade kategoridata givetvis numeriska i den meningen att antalet utfall inom varje kategori räknas samman. Vidare kan numeriska data omvandlas till kategoridata genom att man förutbestämmer ett visst värde som gräns. Man får då två kategorier: överskrider/överskrider inte gränsen. Detta kan kallas att man »dikotomiserar« numeriska data.

Ytterligare en uppdelning av data kan göras. För varje person/patient som accepterar att delta i en undersökning registreras baslinjedata. Därefter registrerar man en eller flera gånger resultatdata (som också kan kallas effektdata, om undersökningen gäller interventioner, dvs behandlande eller förebyggande åtgärder). I båda dessa datagrupper kan såväl kategoridata som numeriska data finnas.

En speciell form av data är de som registreras retrospektivt. Ett viktigt exempel är fall–kontrollundersökning: hos personer som har respektive inte har ett visst medicinskt problem tar man reda på om de varit exponerade för en eller flera faktorer som misstänks ha en orsakande (eller skyddande) roll. Det vanligaste är att både exponeringen och det medicinska problemet registreras som ja/nej, dvs som kategoridata.

Den text som följer gäller huvudsakligen resultatdata av kategorityp från prospektiva undersökningar, dvs kohortstudier eller kliniska prövningar. Sådana resultatdata utgörs oftast av händelser (eng: events). En undersökning kan ha flera resultatvariabler som registreras i form av händelser (som givetvis skall vara noga definierade i prövningsplanen). Händelser kan vara av gynnsam (positiv) eller ogynnsam (negativ) karaktär.

SAMMANFATTAT

- Risktal och riskkvoter är lättare att uppfatta intuitivt än oddstal och oddskvoter.
- Oddskvoter har vissa fördelar och är ibland outhärliga vid analys av data.
- Man måste vara uppmärksam på om kvotall i en resultatredovisning är riskkvoter eller oddskvoter.
- Oddskvoter ger alltid en siffermässigt starkare bild av skillnader och effekter än riskkvoter beräknade från samma data.
- Bara vid låga risktal (högst 10–15 procent) är det acceptabelt att betrakta en oddskvot som en approximation av motsvarande riskkvot.

Evidensbaserad medicin

En händelse benämns i engelsk text ofta »endpoint«. Ordet är lämpligt bara då det är frågan om dödlig utgång eller en händelse som är förutbestämd att innebära att patientens deltagande i studien upphör. I praktiken används det ofta också om andra händelser, och denna breda användning ser man tyvärr även i svenska texter. Ordet »endpoint« är överflödigt i svenskan och kan alltid ersättas av »händelse«.

Risk och riskreduktion

Antalet händelser av visst slag som observerats under en viss tidsrymd används för att i efterhand skatta sannolikheten för att händelsen skulle inträffa. I klinisk forskning fås sannolikheten som antalet patienter som upplevt händelsen dividerat med det totala antalet deltagare, alltså ett andelstal. De händelser som studeras är ofta av oönskat slag, t ex insjuknande, försämring, död. Sannolikheten kan då kallas risk. En risk kan variera mellan 0 och 1 (0–100 procent).

Exempel: Man jämförde läkemedlen A och B beträffande förmågan att förebygga recidivinfarkt hos personer som just

haft hjärtinfarkt. Varje patient i de båda grupperna följdes under en förutbestämd tidsperiod, med undantag av de patienter som fick hjärtinfarkt under perioden och som därefter inte observerades vidare i studien. För enkelhets skull förutsätter vi att alla patienter som inte fick hjärtinfarkt observerades tiden ut, dvs ingen annan händelse än hjärtinfarkt medförde att någon patient måste lämna studien. Risk (risktalet) för hjärtinfarkt i vardera gruppen utgörs då av antalet personer som fick hjärtinfarkt dividerat med totala antalet i gruppen. Resultaten visas i Tabell I.

Skillnaden mellan grupperna kan uttryckas som riskdifferensen: $0,226$ minus $0,132 = 0,094$ (eller $9,4$ procentenheter; observera: inte $9,4$ procent).

Differensen kan i detta sammanhang också kallas riskreduktionen, eftersom den ger ett uttryck för A-läkemedlets förmåga att reducera infarktstrisken från den nivå som observerades hos dem som fick B. Närmare bestämt är detta den absoluta riskreduktionen (ARR).

Om ARR-talet inverteras får man »number needed to treat« (NNT), som här blir $1/0,094 = 10,6$. Talet anger att i genomsnitt elva patienter behöver få läkemedel A för att en av dem skall undgå recidivinfarkt under den tidsperiod som studerades, detta i jämförelse med vad som skulle hänt om patienterna fått läkemedel B.

En differens kan också uttryckas i relativa tal, i detta fall genom att riskdifferensen uttrycks som andel av det högre risktalet. Denna relativa riskreduktion (RRR) blir här $0,094/0,226 = ca\ 0,416$ (eller $41,6$ procent), alltså rent siffermässigt betydligt högre än den absoluta reduktionen.

Den relativa riskreduktionen får naturligtvis inte feltolkas som följer: »behandling A medför att de infarkter som inträffar blir i genomsnitt ca 40 procent lindrigare än vad fallet blir med behandling B«. NNT-begreppet är entydigt: i varje grupp av genomsnittligen elva patienter, som under tidsperioden ifråga får behandling A, kommer en patient att dra fördel av att medlet har bättre effekt än B. För de övriga tio innebär A ingen bevisad fördel jämfört med B.

Det är viktigt att man inte förväxlar de båda riskreduktionsmåten ARR och RRR. När läkemedelstillverkare sprider kännedom om prövningsresultat föredrar de oftast den mera imponerande relativa siffran. Tillförlitlig bedömning av läkemedlets värde kräver att även absoluta siffran redovisas. Men den absoluta riskreduktionen har också sina begränsningar: de siffermässigt likstora differenserna $54-51$ procent och $7-4$ procent är knappast likartade. En korrekt resultatredovisning måste alltså förutom reduktionssiffrorna (differenserna) innefatta grunddata, dvs vardera gruppens risksiffra.

Riskkvot – ett annat sätt att uttrycka olikhet i risk

Ett annat viktigt sätt att uttrycka en olikhet i risk är att bilda en kvot mellan risktal. Man får då en riskkvot (eng: risk ratio) som också kan kallas relativ risk. Förkortningen RR används ofta.

I vårt exempel blir $RR\ 0,132/0,226 = 0,584$.

$RR = 1$ innebär förstås att någon skillnad mellan undersökta grupper inte har påvisats. För RR-värden gäller detsamma som för absoluta och relativa riskreduktioner: en tillförlitlig bedömning av resultatet kräver att man har kännedom om de risktal som kvoten bildats av.

1 minus RR ger den relativa riskreduktionen. I exemplet får vi $1 - 0,584 = 0,416$, samma tal som i beräkningen ovan.

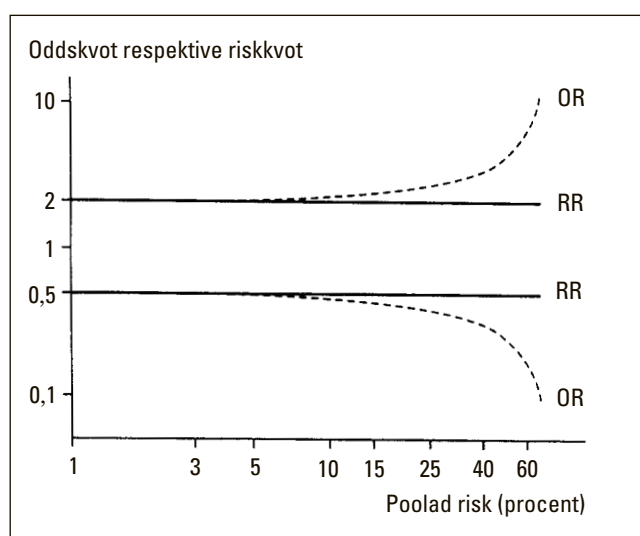
En riskkvot är lätt att tolka. Om en behandlingsprövning ger RR $0,80$ var risken i den behandlade gruppen 80 procent av risken i kontrollgruppen. Annorlunda uttryckt blev den relativa riskreduktionen 20 procent, ett mått på behandlingseffekten.

I undersökningar av exponeringar som medför riskökning

Tabell I. Resultat av läkemedelsprövning avseende prevention av hjärtinfarktrecidiv.

	Antal personer	Antal händelser	Risk	Odds
Läkemedel A	144	19	19/144 (0,132)	19/125 (0,152)
Läkemedel B	146	33	33/146 (0,226)	33/113 (0,292)

ARR	9,4 procentenheter
NNT	11 (under den tidsperiod som studien pågick)
RRR	41,6 procent
RR	0,584
OR	0,521



Figur 1. Relationer mellan riskkvoter (RR) och oddskvoter (OR). Kurvorna visar hur OR progressivt avviker från $RR = 2$ respektive från $RR = 0,5$ när risktalet ökar. X-axeln avser den poolade risken, dvs det risktal som erhålls då de två gruppernas incidensstal och deltagarantal slås samman. Diagrammet är med medgivande hämtat från *Systematic Reviews in Health Care: Meta-analysis in context* (Egger M, Davey Smith G, Altman DG, editors), 2nd ed. London: BMJ Books; 2001, p. 31.

blir RR över 1 . Om RR är 3 var den exponerade gruppens risk tre gånger högre än den icke-exponerade gruppens. Detta kan också formuleras så att exponeringen höjde risken med 200 procent.

En nackdel med termen »risk« är att undersökningar kan gälla händelser som är önskvärda. Man kan t ex registrera hur stor andel av en grupp patienter som klarar att gå en viss distans efter ett visst ortopediskt ingrepp, eller hur stor andel av en grupp kvinnor som blir gravida inom en viss tidsperiod när de får ovulationsstimulerande behandling. Man kan då, och även när det gäller negativa händelser, använda termen incidens (eng: event rate). Den betyder i det här sammanhanget antalet patienter hos vilka händelsen i fråga observeras dividerat med det totala antalet patienter i undersökningen. Liksom med termen risk kan man naturligtvis beräkna deriverade data: absolut och relativ incidensdifferens samt incidenskvot.

Hazard inte synonymt med risk

I allmänspråket kan »hazard« översättas med »risk«, men i statistiken är begreppen inte synonyma. Som statistisk term

står hazard (termen används också på svenska) för momentan risk, dvs sannolikheten för att en viss händelse skall inträffa vid en viss tidpunkt (litet tidsintervall), givet att den inte inträffat tidigare. Hazardbegreppet är viktigt i vissa statistiska metoder som inte bara analyserar skillnader i antalet händelser utan också beaktar skillnader i vilken tid som förflutit till dess en händelse inträffat (eng: time to event). Relationen mellan två hazardtal anges oftast med en hazardkvot.

Odds kan variera från 0 till oändligheten

Ett odds (oddstal) är sannolikheten för att en viss händelse skall inträffa dividerad med sannolikheten för att den inte skall inträffa. När man drar ett kort ur en kortlek är sannolikheten för att det skall vara ett spaderkort $1/4$. Odds är däremot $1/3$, vilket utläses »ett mot tre«.

För sannolikhet används ofta notationen P (för probabilitet). Sannolikheten för händelsen H betecknas P(H). Sannolikheten för att H inte skall inträffa är $1 - P(H)$.

$$\text{Odds blir alltså } \frac{P(H)}{1 - P(H)}$$

Ett odds kan variera från 0 till oändligheten. Ett sannolikhetstal kan enligt ovan lätt räknas om till ett odds. Omvänt kan ett odds räknas om till en sannolikhet:

$$\text{Sannolikhet} = \frac{\text{odds}}{1 + \text{odds}}$$

Odds kontra risk

Inom medicinen är risktal mera bekanta och lättolkade än oddstal. Om man vid en hjärtinfarktenhet har en mortalitet under vårdtiden av 14 procent är detta ett risktal som också kan skrivas 0,14. Det ligger inte så nära till hands att bedöma behandlingsresultatet genom att beräkna oddstalet, som i det här fallet blir $14/86 = \text{ca } 0,163$ eller drygt 16 procent.

Odstal och risktal kommer närmare varandra ju lägre risken är. Vid höga risker blir de mycket olika; exempelvis svarar risk 0,5 mot odds 1.

Odstskvot kontra riskkvot

En jämförelse av två oddstal (t ex i två patientgrupper) görs oftast genom att man bildar en oddskvot. Som förkortning används i regel OR (eng: odds ratio).

Hur förhåller sig OR till riskkvoten (RR) från samma observationer? Om OR är 1 är RR också 1. I alla andra fall medför den matematiska skillnaden mellan odds och risk att OR och RR skiljer sig åt. Vid alla OR som är mindre än 1 är motsvarande RR-tal högre. När OR är över 1 är motsvarande RR lägre. I samtliga fall visar OR-talet därför en större differens mellan grupperna än RR.

Alltså: om man tror att ett redovisat OR-tal är ett RR-tal får man alltid en överdriven uppfattning om den konstaterade differensen. Detta bedömningsfel blir större ju högre risktalen är. OR börjar avvika märkbart från RR när de jämförda gruppernas poolade risktal är 10–15 procent. Se Figur 1.

Skillnaden RR – OR kan illustreras med hjälp av data från studien av prevention av hjärtinfarkt (Tabell I).

$$\begin{aligned} \text{Riskkvoten blir } & 0,132/0,226 = 0,584. \\ \text{Odstskvoten blir } & 0,152/0,292 = 0,521. \end{aligned}$$

OR är, som alltid vid OR-värden under 1, lägre än RR. OR ger därför en något mera positiv bild av A-läkemedlets effekt, men skillnaden är inte dramatisk.

Man kan inte direktomvandla ett OR-tal till motsvarande RR-tal (eller omvänt), eftersom relationen mellan OR och RR påverkas av risktalens nivå. Fakta 1 visar omräkningsprocedurer för risk, odds och kvoter.

II Fakta 1

Omvandling riskkvot (RR)–oddskvot (OR)

Omräkning av ett RR-tal till motsvarande OR-tal, eller omvänt, kan göras på två sätt.

- *Genom omvandling av risk till odds respektive av odds till risk*

Man använder formlerna

$$\text{odds} = \frac{\text{risk}}{1 - \text{risk}} \quad \text{risk} = \frac{\text{odds}}{1 + \text{odds}}$$

I vårt exempel var risktalen 0,132 respektive 0,226.

Riskkvoten RR = 0,584.

Omräkning ger oddstalen 0,152 respektive 0,292.

Odstskvoten OR = 0,521.

Belysande jämförelse:

Risktalen 0,53 respektive 0,90 ger nästan samma riskkvot RR som ovan: 0,589.

I detta fall blir oddstalen 1,13 respektive 9, och oddskvoten OR blir 0,126.

RR och OR avviker starkt från varandra vid dessa höga risktal.

- *Med omräkningsformel*

Med följande formler krävs ingen omräkning risk–odds eller omvänt

$$\text{OR} = \frac{\text{RR} \times (1 - R_k)}{1 - R_k \times \text{RR}} \quad \text{RR} = \frac{\text{OR}}{1 - R_k \times (1 - \text{OR})}$$

där R_k står för det högsta av de båda risktalen (kontrollgruppens risktal i en behandlings- eller profylaxprövning).

Analogt med risk respektive odds gäller att det är lättare att uppfatta den numeriska innebörden av en riskkvot än av en oddskvot. Om RR är 1,34 ser man genast att risken (incidensen) är 34 procent högre i den ena gruppen än i den andra. Om siffran i stället anger en oddskvot är det svårare att göra en sådan direkt numerisk tolkning.

Sammanfattningsvis kan en oddskvot användas som en approximation av motsvarande riskkvot, men bara vid låga risktal, se ovan. Det är oacceptabelt att (som ibland sker) betrakta begreppen som i praktiken identiska. Det förekommer till och med att publicerade resultatstabeller saknar uppgift om vilken av de båda kvoterna som siffrorna gäller.

Fördelar med odds och oddskvoter

Trots de beskrivna olägenheterna med odds och oddskvoter förekommer de ofta i medicinsk litteratur. Detta beror bl a på följande:

– Fallkontrollundersökningar: data kan inte ge någon annan information än odds och oddskvoter. Metoden för insamling av data medför nämligen att man inte får några risktal (incidenstal) för vare sig sjukdom eller exponering. Man kan alltså inte beräkna riskkvoter.

– Bayesiansk analys vid diagnostik med hjälp av det diagnostiska testets likelihoodkvoter (LR) bygger på odds som mått på diagnosens sannolikhet. Odds före test multipliceras med testets likelihoodkvot LR+ för positivt testutfall, respektive med LR– för negativt testutfall, för att ge odds efter test (som kan omvandlas till en sannolikhet). För att slippa omräk-

andet kan man använda det sk Fagan-nomogrammet, uppkallat efter sin uppfinnare [1]. Ett exempel ges i en artikel som beskriver bayesianskt tänkande vid diagnostik [2].

– Risk- och oddstal, och deras respektive kvoter, skiljer sig åt i matematiska egenskaper. Dessa skillnader är för det mesta till fördel för odds och oddskvot. Några exempel följer:

Ett risktal kan variera mellan 0 och 1. Även riskkvoter har begränsad variation uppåt. Om man undersöker omständigheter som höjer en risk är den maximala riskkvoten begränsad av »basrisken«. Om denna t ex är 0,4 kan riskkvoten högst bli 2,5 (risken kan högst bli 1, som divideras med 0,4).

Odds och oddskvoter kan variera från 0 till oändligheten. Logaritmerade odds och oddskvoter varierar obegränsat kring noll, som motsvarar »neutralläget« där odds respektive oddskvot är 1. Logaritmeringen medför också att ett visst tal (odds eller oddskvot) och dess inverterade värde ligger på var sin sida om, och med samma avstånd till, nollpunkten.

Ett oddstal är meningsfullt även om det inverteras: om odds för dödlig utgång är 0,20 får man genom invertering talet 5 som är odds för överlevnad. Om det i stället är risken för dödlig utgång som är 0,20 säger det inverterade talet att det totala antalet patienter var fem gånger större än antalet avlidna patienter – en information som knappast uppfattas som relevant.

Motsvarande gäller för oddskvoter: om man utför resultat-kalkylen i båda riktningarna (t ex med utgångspunkt från överlevnad respektive dödlighet) blir de båda oddskvoterna varandras inverterade värden. Detta gäller inte för riskkvoter. Om t ex två behandlingar jämförs med hjälp av incidenskvoter för »symtomfrihet« respektive för »kvarvarande symtom« blir resultaten inte varandras inverterade värden. Utfallet kan till och med bli en statistiskt signifikant skillnad i ena fallet men icke-signifikant skillnad i andra fallet, fastän grunddata är desamma.

– Oddskvoternas egenskaper är fördelaktiga vid vissa statistiska kalkyler, bl a sådana som används när resultat av olika undersökningar skall slås samman och analyseras i en metaanalys. Tabeller och figurer i publicerade metaanalyser använder därför ofta oddskvoter i resultatredovisningen. Ibland redovisas också resultat av individuella undersökningar med oddskvoter. Orsaken kan vara att författarna använt en analysmetod som bygger på oddstal. Läsare som föredrar riskkvoter kan räkna om publicerade oddskvoter till riskkvoter med någon av metoderna i Faktarutan, förutsatt att endera oddstalen eller kontrollgruppens risktal är redovisade.

Statistisk bedömning

Den här artikeln har nämnt flera olika sammanfattande relationstal som kan användas för att beskriva resultat av undersökningar med kategoridata som resultatmått: riskdifferens, absolut och relativ riskreduktion, NNT (number needed to treat), riskkvot och oddskvot. Varje sådan resultatangivelse måste givetvis analyseras statistiskt så att man kan bedöma sannolikheten för att resultatet uppstått som följd av slumpens verk. Här är inte platsen att redogöra för alla tänkbara metoder, men konfidensintervall är en välkänd och lättförståelig möjlighet. En utmärkt bok med tillhörande programdiskett [3] är helt ägnad åt beräkning av konfidensintervaller.

Bland övrig litteratur inom området kan två verk rekommenderas: ett standardverk inom klinisk epidemiologi [4] och en bok som helt är ägnad metoder för systematiska översikter och metaanalyser [5].

Sammanfattning

Vetenskapliga resultat som utgörs av kategoridata i form av antal händelser beskrivs ofta i form av risk- eller oddstal. Om ett visst resultat uttrycks som både risk och odds blir dessa tal

aldrig identiska. Oddstalet är alltid högre. Skillnaden blir större ju högre risktalet är.

Jämförelse mellan två risktal kan framför allt i behandlingsprövningar göras genom att man beräknar absolut riskreduktion (ARR) eller relativ riskreduktion (RRR). Det inverterade värdet av ARR ger »number needed to treat« (NNT), det antal patienter som behöver få den prövade behandlingen under angiven tid för att en av dem skall undgå den studerade händelsen. Relationen mellan två risktal kan också uttryckas som kvoten mellan dem, riskkvot (relativ risk, RR).

Jämförelse mellan två oddstal görs nästan alltid genom att man bildar en oddskvot (odds ratio, OR).

Om ett visst resultat uttrycks som både riskkvot och oddskvot är dessa lika endast då talen är 1. I samtliga andra fall ger oddskvoten ett siffermässigt större utslag, dvs visar jämfört med riskkvoten en mera markant grupp-differens. Skillnaden mellan oddskvot och motsvarande riskkvot blir större ju högre risktalen är.

Det är lättare att uppfatta den kvantitativa innebörden av risktal och riskkvoter än av oddstal och oddskvoter. Oddskvoter har gynnsamma matematiska egenskaper, vilket medför att de ofta används i bl a metaanalyser. När man tar del av publicerade resultat måste man vara uppmärksam på om angivna kvotsiffror utgör riskkvoter eller oddskvoter.

Referenser

1. Fagan TJ. Nomogram for Bayes' theorem. *New Engl J Med* 1975; 293:257.
2. Taube A, Malmquist J. Räkna med vad du tror. Bayes' sats i diagnostiken. *Läkartidningen* 2001;98:2910-3.
3. Altman DG, Machin D, Bryant T, Gardner MJ, editors. *Statistics with Confidence*. 2nd ed. London: BMJ Books; 2000.
4. Fletcher RH, Fletcher AW, Wagner EH. *Clinical Epidemiology. The essentials*. 3rd ed. Baltimore, MD: Williams & Wilkins; 1996.
5. Egger M, Davey Smith G, Altman DG. *Systematic Reviews in Health Care. Meta-analysis in context*. 2nd ed. London: BMJ Books; 2001.

SUMMARY

Risk and odds – how to calculate with events

Jörgen Malmquist

Läkartidningen 2002;99:751-6

Categorical scientific data consisting of counts of events are frequently reported as risk or odds figures. Given a certain set of data, odds always differs upwards from risk. The relation between two risk figures may be expressed in various ways, one of which is the risk ratio. In the case of two odds figures, the choice is almost always the odds ratio. With a certain set of data from two groups, the odds ratio is not identical to the risk ratio (except when both are equal to 1). The odds ratio always magnifies the intergroup difference. When assessing published data, one must take care to observe whether reported ratio figures denote risk ratios or odds ratios.

Correspondence: Jörgen Malmquist, V Tallstigen 6, SE-236 42 Höllviken, Sweden. (jorgen.malmquist@mailbox.swipnet.se)

Surrogatmått – en ersättning för det man egentligen vill mäta

Lättmätta men irrelevanta?

■ Ett rimligt krav på en metod som skall användas i klinisk praxis är att det finns dokumentation som visar att användning av metoden för patienten gör större nytta än skada, eller, med andra ord, att balansen mellan nytta och risk är gynnsam. Detta gäller inte bara behandling (t ex med läkemedel), kirurgiska ingrepp, sjukgymnastik, psykoterapi och metoder inom alternativmedicinen, utan också diagnostik (t ex screening för tidig upptäckt av bröst- eller prostatacancer).

När det gäller läkemedel anger Läkemedelslagen att det åligger tillverkaren att visa »ändamålsenlighet« för att erhålla godkännande för marknadsföring: »Ett läkemedel är ändamålsenligt om det är verksamt för sitt ändamål och vid normal användning inte har skadeverkningar som står i missförhållande till den avsedda effekten.« Det som avgör ett läkemedels ändamålsenlighet är således balansen mellan nytta och risk.

För andra behandlingsmetoder och för diagnostiska åtgärder är kraven inte lika väldefinierade; det kan räcka med att några inom vården tycker att det förefaller spännande och rimligt att pröva en ny metod, t ex laparoskopisk kirurgi, för att metoden skall kunna komma till användning i klinisk praxis.

De flesta exempel som diskuteras i föreliggande artikel avser behandling med läkemedel, men samma principer kan tillämpas när det gäller andra interventioner.

Vad är ett surrogatmått?

Det engelska uttrycket för surrogatmått är »surrogate endpoint« eller »surrogate marker«. Begreppet har av Robert Temple vid amerikanska Food and Drug Administration (FDA) definierats på följande sätt: »A surrogate endpoint of a clinical trial is a laboratory measurement or a physical sign used as a substitute for a clinically useful endpoint that measures directly how a patient feels, functions or survives. Changes induced by a therapy on a surrogate endpoint are expected to reflect changes in a clinically meaningful endpoint« [1].

Den ordlista som sammanställts av Statens beredning för medicinsk utvärdering (SBU) ger följande definition:

»Surrogatmått är ett effektmått som har ett samband med den effekt som är det egentliga slutmålet för en åtgärd och som i en klinisk prövning används som ersättare för detta.«

Vad är det man egentligen vill mäta?

Det vanligaste sättet på vilket en behandling kan göra nytta är att åtgärden får patienten att må bättre, t ex genom lindring av

SAMMANFATTAT

Värdet av en behandlingsmetod bestäms av balansen mellan de positiva och de negativa effekterna. Att dokumentera nytta kräver ofta att stora grupper av patienter följs över lång tid, vilket är dyrt och ofta förenat med etiska, praktiska och metodologiska problem.

Att indikera nytta genom att i stället mäta något som är enkelt att mäta kan förefalla tilltalande. Ett surrogatmått är ett laboratorievärde eller ett fysikaliskt tecken som används som ersättning för det man egentligen vill mäta. Förhoppningen är att förändringar i surrogatmättet skall avspeglar förändringar i det man egentligen ville mäta.

Den stora fördelen med ett väl fungerande surrogatmått är att man med färre patienter och på betydligt kortare tid kan dokumentera nyttan av en behandling.

Dokumentationen är dock bristfällig för att förändringar av surrogatmättet verkligen avspeglar något kliniskt betydelsefullt.

Evidensbaserad medicin

symtom som smärta, feber, ångest, andnöd eller nästäppa. »Må bättre« kan också innebära ett ökat allmänt välbefinnande eller en förbättrad hälsorelaterad livskvalitet.

En behandling kan också göra nytta genom att minska risken såväl för framtida icke-dödliga komplikationer till sjukdomen som för progress av funktionsnedsättning. Det kan t ex handla om att reducera risken för exacerbationer vid astma, för synnedsättning vid diabetes, för kotfraktur vid osteoporos eller för icke-letal hjärtinfarkt vid hyperkolesterolemi. Progress av funktionsnedsättning, t ex lägre expiratoriskt flöde vid astma, får anses vara ett surrogatmått så länge patienten inte märker av nedsättningen [2]. Detsamma gäller progress

av njurfunktionsinskränkning vid hypertoni. Först när patienten kan notera den försämrade funktionen kan man tala om ett egentligt patientrelaterat effektmått.

Ett tredje sätt på vilket en behandling kan göra nytta är genom att förlänga patientens liv, t ex vid cancer, aids eller efter hjärtinfarkt.

Varför mäter man inte alltid det man egentligen vill mäta?

Det kan tyckas rimligt att den som vill introducera en ny behandlingsmetod åläggs att dokumentera nyttan av densamma i form av gynnsam effekt enligt något av ovanstående kriterier. Problemet är att det ofta finns metodologiska svårigheter, tar lång tid, kräver ett stort antal patienter i kliniska prövningar, är dyrt eller är tvivelaktigt ur etisk synvinkel. Det kan därför vara lockande att i stället mäta ett surrogatmått, eftersom detta oftast är enklare, inte tar lika lång tid, kräver färre patienter, kostar avsevärt mindre och kan förefalla etiskt oantastligt.

Det sistnämnda kan dock ifrågasättas, eftersom Helsingforsdeklarationen kräver att klinisk forskning både skall vara meningsfull och ge kunskap till nytta för patienter med den sjukdom som studeras.

Svårigheter vid dokumentation av symtomlindring

Vid icke livshotande tillstånd är den vanligaste orsaken till behandling att man vill att patienten skall må bättre, med lindrigare symtom eller ökad hälsorelaterad livskvalitet, helst bådadera.

Mätning av subjektiva symtom vid t ex angina pectoris eller refluxsjukdom medför flera metodologiska problem. Patienter med angina pectoris kan beskriva sina symtom på olika sätt: som »smärta«, »sveda«, »trånghets känsla« eller »tyngdkänsla«. Refluxpatienten kan vara mest besvärad av »halsbränna« eller »sveda bakom bröstbenet«. Man kan således stöta på problem redan vid benämning av de viktigaste symtomen.

Ett annat problem vid mätning av symtom är att patienter med samma diagnos inte alltid har identiska problem. Man räknar med att ca 40 procent av astmapatienter får nattsömnen störd av sin lungsjukdom. För denna grupp är minskad nattlig astma en betydande behandlingsvinst. För de 60 procent som inte lider av nattliga symtom finns inte denna behandlingsvinst att hämta. Om man i en klinisk prövning blandar patienter

med och utan nattlig astma, och har lindring av nattlig astma som ett viktigt mått för nytta, är risken stor att positiva resultat för en patientgrupp späds ut av det faktum att en större grupp inte har något att vinna när det gäller detta symtom.

Ytterligare ett problem vid mätning av subjektiva symtom är mätskalan. Ofta används en sk ordinalskala (t ex ingen–lätt–måttlig–svår–mycket svår), där man vet ordningsföljden på skalstegen men inte om skalstegen är lika höga, vilket medför betydande begränsningar i den statistiska bearbetningen. Vid t ex smärta används inte sällan en visuell, analog skala (VAS), en 100 mm lång linje. Svårigheterna med denna är dels att veta hur kliniskt relevant en skillnad på t ex 5 mm är, dels att det kan ifrågasättas om en förändring på 1 mm i mitten av skalan representerar en lika stor förändring i smärta som 1 mm i den nedre eller den övre delen av skalan.

Sammantaget kan man konstatera att det finns flera metodologiska problem när man vill dokumentera att en behandling leder till symtomlindring. Det kan därför vara lockande att i stället använda något instrument för att få fram ett mätvärde som lämpar sig väl för statistisk bearbetning. Vid astma kan man välja att i stället för symtom mäta den forcerade exspiratoriska volymen under en sekund (FEV_{1s}), eller under 10 millisekunder, »peak expiratory flow« (PEF). Vid refluxsjukdom förlitar man sig ofta mer på gastroenterologens subjektiva bedömning av slemhinnerodnad än på patientens bedömning av sina symtom.

Hur dokumentera ökad hälsorelaterad livskvalitet?

Hälsorelaterad livskvalitet är ett informativt effektmått i kliniska studier [3]. Inom flera terapiområden finns sjukdomspecifika livskvalitetsmått med känd reliabilitet, validitet och känslighet för kliniskt relevanta förändringar. Viktiga fördelar med livskvalitet som effektmått är att mätningen är patientbaserad, mångdimensionell och ger en bild av hur patienten mår och fungerar i vardagen.

Kliniska prövningar är en blandning av klinik och statistik. Metodiken vid kliniska prövningar bygger på att testa i förväg uppställda hypoteser. Det innebär att man i prövningsprotokollet skall definiera en eller ett fåtal frågeställningar, vilken mätmetod som skall användas och när mätning skall ske. Formulär för livskvalitetsmätning innehåller ofta ett stort antal frågor, vilket är väl motiverat för att möjliggöra en allsidig beskrivning av hur patienten mår. Problemet är att om man vill signifikantesta för varje enskild fråga är risken för sk slumpsignifikanser stor. Om man i stället slår ihop närliggande frågor riskerar man att förlora viktig information. Ytterligare ett alternativ kan vara att i förväg välja ut en av frågorna som primär frågeställning. Då kommer man ifrån problemet med multipla signifikantest, men riskerar i stället de problem som ovan beskrivits vid mätning av enskilda symtom.

Mätning av livskvalitet i kliniska prövningar tillför utan tvivel värdefull information om patienternas allmänna välbefinnande. Ändå ser man relativt sällan att mätning av livskvalitet är primär frågeställning. Den främsta orsaken till detta är den inbyggda konflikten mellan antalet frågor i livskvalitetsformulär och de statistiska metoderna för hypotesprövning, vilka utgör basen för utvärdering av resultat av kliniska prövningar.

Hur mäta risk för komplikationer och funktionsnedsättning?

Många behandlingar ges inte primärt för att öka patientens aktuella välbefinnande utan för att förebygga framtida komplikationer, recidiv eller progress av sjukdomen. Flera olika behandlingar har visat sig minska risken för icke dödliga komplikationer under de första åren efter hjärtinfarkt. Många behandlingar inom onkologin ökar signifikant tiden till recidiv. Vid astma ger behandling med inhalationssteroider långsam-

mare progress av lungfunktionsnedsättning. Gemensamt för dessa exempel är att komplikationsincidensen är relativt hög inom något eller några år. Det är därför möjligt att visa nytta genom kliniska prövningar som pågår under ett fåtal år och omfattar upp till några tusen patienter.

Svårigheterna är betydligt större om man vill visa att blodtryckssänkande behandling vid mild-måttlig hypertoni hos medelålders individer minskar risken för slaganfall, hjärtinfarkt eller hjärtsvikt. En klinisk prövning för att fastställa detta skulle kräva en mycket stor, långvarig och dyr prövning. Man kan förstå de företag som hellre väljer att i en tolvveckorsstudie visa att blodtrycket går ner. Vid typ II-diabetes är det betydligt enklare att dokumentera korttidseffekter på blodsocker eller HbA_{1c} än att visa att tiden till synnedsättning, neuropati eller njursvikt förlängs.

Hur dokumentera förlängd överlevnad?

Nyttan kan knappast ifrågasättas av en behandling som visats ge förlängd överlevnad. Den kliniska relevansen är ju odiskutabel, och mätmetoden ger inte mycket utrymme för olika tolkningar. Detta gäller studier med signifikant effekt på totalmortalitet, men knappast prövningar som visar t ex reducerad kardiovaskulär mortalitet men ingen skillnad i total dödlighet.

Svårigheterna förknippade med att dokumentera att en diagnostisk åtgärd eller behandling leder till förlängd överlevnad hänger förstas samman med prognosen vid den studerade sjukdomen. För sjukdomar som subaraknoidalblödning eller pankreascancer kan det vara möjligt att med ett rimligt patientantal och på relativt kort tid visa att en ny behandlingsmetod kan ge förlängd överlevnad. För andra sjukdomar är det svårt att med kliniska prövningar visa att en viss åtgärd förlänger livet. Det gäller även för sjukdomar som är vanliga och som är förenade med hög mortalitet.

Ett aktuellt och omdiskuterat exempel är mammografiscreening och dess effekt på totalmortaliteten. Vid astma finns det indikationer från observationella studier att behandling med inhalationssteroider förbättrar överlevnaden, men detta går knappast att studera i en prospektiv klinisk prövning.

Vilka är fördelarna med surrogatmått?

Ett surrogatmått som predicerar nytta kan bidra till att nya effektiva behandlingar blir tillgängliga för sjukvården flera år innan dokumentationen för verklig klinisk nytta blir tillgänglig. Ett exempel på detta är simvastatin, som godkändes för marknadsföring på basis av mätning av lipidnivåer och där senare kliniska prövningar har påvisat nytta vad gäller kardiovaskulär morbiditet och mortalitet.

Kostnaden för kliniska prövningar med effekt på ett surrogatmått som primär frågeställning är avsevärt lägre, vilket förhoppningsvis kan ge utslag vid prissättningen av läkemedlet.

... och nackdelarna?

Effekt på ett surrogatmått har i sig inte något värde för patienten. Ofta väljs surrogatmått baserat på logiskt tänkande. En riskfaktor, t ex blodtryck eller serumkolesterol, samvarierar med morbiditet och mortalitet. En intervention reducerar riskfaktornivån, och därför är det rimligt att förmoda att interventionen minskar risken för morbiditet och mortalitet.

En anledning till att en viss reduktion av en riskfaktor inte behöver ge motsvarande reduktion i morbiditet och mortalitet är att riskfaktorn kan vara en varningssignal för andra rubbningar i kroppen. Det förhöjda blodtrycket kan vara en del av ett metabolt syndrom med störningar i t ex insulinresistens och lipidmetabolism. Ett blodtryckssänkande preparat kan då »skruva ur varningslampan«, dvs sänka blodtrycket men en-

dast obetydligt minska risken för morbiditet och mortalitet genom att andra riskfaktorer förbises.

En annan anledning till att en gynnsam effekt på ett surrogatmått inte behöver vara gynnsam för patienten är att alla läkemedel även har negativa effekter som kan påverka balansen mellan nytta och risk. Surrogatmålet ger ingen information om riskerna. Ett exempel är Atromidol (klofibrat), under 1960- och 1970-talen en storsäljare vid hyperkolesterolemi. Klofibrat har en gynnsam effekt på lipidnivåerna och borde logiskt sett kunna ge minskad morbiditet och mortalitet. En stor europeisk placebokontrollerad klinisk prövning under 1970-talet visade överraskande en signifikant ökad mortalitet i klofibratgruppen. Ökningen gällde även det som behandlingen skulle minska risken för, nämligen kardiovaskulär mortalitet [4]. Gynnsam effekt på ett surrogatmått behöver alltså inte betyda minskad risk för patienten.

Ytterligare en orsak till varför det finns brister i sambandet mellan effekter på surrogatmått och effekter på morbiditet/mortalitet är att läkemedel kan ha flera positiva effekter. Detta verkar vara fallet för vissa av statinerna. Utöver gynnsamma effekter på lipidmönstret har de visats ha flera andra biologiska effekter, bl a antiinflammatoriska och plackstabiliserande. Graden av dessa effekter tycks inte vara kopplad till graden av sänkning av LDL-kolesterol. I den nyligen rapporterade Heart Protection Study hade simvastatin en gynnsam effekt på morbiditet och mortalitet oberoende av om kolesterolnivån före studien var förhöjd, normal eller låg under gängse behandlingsmål [5]. Det förefaller rimligt att behandla patienter med ökad kardiovaskulär risk oavsett lipidnivåer, och i stället för att behandla laboratorievärden välja preparat och dos på grundval av vad som i kliniska prövningar visat sig påverka morbiditet och mortalitet.

Vilka förändringar avspeglar surrogatmålet?

Ett mycket stort antal kliniska prövningar med olika typer av läkemedel har visat att sänkning av blodtrycket vid hypertoni minskar risken för slaganfall. Detta har medfört att det råder enighet om att sänkning av ett förhöjt blodtryck är en tillförlitlig markör för minskad risk för slaganfall. Detta betyder inte att det råder enighet om blodtryckssänkning som surrogatmått för nytta vad gäller andra komplikationer, t ex hjärtinfarkt och hjärtsvikt. I den pågående ALLHAT-studien avbröts doxazosinarmen i förtid av etiska skäl, då antalet kardiovaskulära komplikationer i denna grupp var 25 procent större än i klortalidongruppen, trots likvärdig blodtryckssänkning [6].

Signifikant blodtryckssänkning i en tolv veckor lång klinisk prövning av ett nytt blodtryckssänkande läkemedel indikerar nytta i form av reducerad risk för slaganfall. Prövningen ger dock föga vägledning om balansen mellan nytta och risk. Dels ger den inte information om andra tänkbara positiva effekter i form av reducerad risk för hjärtinfarkt, hjärtsvikt eller njurskada, dels ger den knappast någon vägledning för bedömning av riskerna vid långtidsbehandling.

I kliniska prövningar vid aids användes tidigt antalet CD4-celler som surrogatmått för nytta, definierat som förlängd tid till insjuknande i aids för HIV-positiva eller förlängd överlevnad för aidssjuka. I en översikt 1994 [7] visade det sig att antalet CD4-celler ökade signifikant i sju av åtta studier där behandlingen förbättrade prognosen. Problemet var att CD4-cellerna ökade signifikant även i sex av åtta studier där behandlingen inte förbättrade prognosen.

Senare har man valt att använda virusförekomst i blod, »viral load«, som surrogat för nytta. På basis av detta effektmått kom ett flertal HIV-medel snabbt ut på marknaden, och för några av dessa har klinisk nytta senare visats i flera studier. Virusförekomst kan dock inte i alla lägen förutsäga klinisk ef-

fekt. Goda kliniska resultat kan ses även om man inte helt lyckas pressa ned virusreplikationen [8].

Det finns misslyckanden

Kartan överensstämmer dock inte alltid med verkligheten. Listan över misslyckanden med surrogatmått kan dessvärre göras lång. Här är två exempel:

- *Patienter med kombinationen koronarsjukdom–ventrikulära extrasystolier (ES) utgör en högriskgrupp för plötslig död.* För drygt tio år sedan introducerades i USA två nya antiarytmika, som mycket effektivt reducerade antalet ES enligt långtidsregistrering av EKG. Rent logiskt borde dessa läkemedel kunna minska risken för plötslig död. Logiken ledde till att mer än 200 000 patienter per år behandlades med dessa läkemedel i USA.

Efter ingående diskussioner om det etiska i att genomföra en prövning för att studera mortalitet accepterades en placebokontrollerad studie med plötslig död som primär frågeställning. Studien avbröts i förtid av säkerhetskommittén sedan det visat sig att 33 fall av plötslig död inträffat med aktiv behandling, mot nio fall med placebo. Totalt avled 56 patienter med aktiv behandling, mot 22 i placebogrupperna [9]. Man räknar med att fler än 50 000 amerikaner dog till följd av behandling med enkainid eller flekainid, vilket är ungefär lika många som dog i Vietnamkriget. Detta är ett drastiskt exempel på riskerna förknippade med att behandla patienter på basis av »mekanismer« eller på vad som förefaller vara logiskt rimligt.

- *Minskad bentäthet är väl korrelerad till risk för frakturer.* Det är därför logiskt att anta att läkemedel som ökar bentätheten minskar frakturrisken. En studie vid Mayo-kliniken visade att natriumfluorid ger en signifikant ökning av bentätheten. Uppmuntrade av dessa resultat startade forskarna en klinisk prövning avseende postmenopausala kvinnor med nedsatt bentäthet. Den nya studien bekräftade att bentätheten ökade, men visade överraskande en signifikant ökad risk för frakturer [10]. Natriumfluorid gav uppbyggnad av ett sköra skelett, som lättare frakturerade.

Konklusion

En fördel med tillförlitliga surrogatmått är att de kan bidra till att effektiva behandlingar snabbare och till lägre kostnader kan komma patienterna till nytta. Exempel på detta finns framför allt inom HIV-området. Vid utveckling av nya behandlingar kan surrogatmått i tidiga studier ge vägledning om huruvida det är värt att gå vidare eller inte. För att verkligen dokumentera nytta krävs dock att patientrelaterade effektmått studeras senare under produktutvecklingen. Det finns anledning att vara skeptisk till läkemedel som långt efter introduktionen på marknaden inte har någon dokumenterad klinisk nytta utan bara har visats påverka laboratorievärden eller fysikaliska tecken.

Ett av problemen med surrogatmått är att de flesta inte är validerade, dvs det är inte visat att en förändring av markören förutsäger en viss klinisk effekt. Ett större problem är att surrogatmåtten inte heller ger information om eventuella risker med en behandling. Surrogatmåtten kan därför bara ge en del av den information som behövs för att bedöma balansen mellan nytta och risk med en intervention.

Referenser

1. Temple RJ. A regulatory authority's opinion about surrogate endpoints. In: Nimmo WS, Tucker GT, editors. Clinical measurement in drug evaluation. New York: J Wiley; 1995.
2. Behandling av astma och KOL. En systematisk kunskapsman-

ställning. Stockholm: SBU, Statens beredning för medicinsk utvärdering; 2000. Rapport nr 151.

3. Sullivan M, Tunsäter A. Hälsorelaterad livskvalitet informativt effektmått i kliniska prövningar. Tillämpningsexempel: astma/kroniskt obstruktiv lungsjukdom. Läkartidningen 2001;98:4428-33.
4. WHO cooperative trial on primary prevention of ischemic heart disease using clofibrate to lower serum cholesterol: mortality follow-up. Report of the Committee of Principal Investigators. Lancet 1980; ii:379-85.
5. Collins R. Heart Protection Study. Föredrag vid 74th annual scientific sessions of the American Heart Association, Anaheim, California November 13 2001.
6. The ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major cardiovascular events in hypertensive patients randomised to doxazosin vs chlorthalidone. The Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial. JAMA 2000;283:1967-75.
7. Fleming TR. Surrogate markers in aids and cancer trials. Stat Med 1994;13:1423-35.
8. Carosi G, Castelli F, Suter F, Maggiolo F, Orani AM, Pan A, et al. Antiviral potency of HAART regimens and clinical success are not strictly coupled in real life conditions: Evidence from the MASTER-1 study. HIV Clinical Trials. 2001;2:399-407.
9. Echt DS, Liebson PR, Mitchell LB, Peters RW, Obias-Manno D, Barker AH, et al. Mortality and morbidity in patients receiving encainide, flecainide or placebo. N Engl J Med 1991;324:781-8.
10. Riggs BL, Hodgson SF, O'Fallon WM, Chao EY, Wahner HW, Muhs JM, et al. Effect of fluoride treatment on the fracture rate in postmenopausal women with osteoporosis. N Engl J Med 1990;322:802-9.

SUMMARY

Surrogate markers –
easy to measure but irrelevant?

Bengt Furberg

Läkartidningen 2002;99:1672-5

Many treatments aim at making a patient feel better due to symptom relief or improved quality of life. Other treatments are given in order to prevent future complications, non-fatal or fatal. There are methodological difficulties in assessing subjective symptoms and quality of life. To document reduced risk of complications takes large, long-term, costly clinical trials. Thus, the use of a surrogate marker to document benefit may seem appealing. A well-validated surrogate marker can demonstrate benefit of a given treatment with fewer patients in a shorter time and at considerably lower cost. It is much easier to show in a 12-week study that a new compound lowers blood pressure than to show that it lowers the risk of stroke, myocardial infarction or renal impairment. One problem with surrogate markers is that very few are well-validated, another is that even those which are cannot be expected to reflect adverse effects, expected or unexpected, and are thus of limited value in the global assessment of benefit vs. risk. It is recommended that results obtained with surrogate markers be regarded as preliminary until results of large trials with clinically relevant outcomes are available.

Correspondence: Bengt Furberg, Krassevågen 7, SE-434 46 Kungsbacka, Sweden (bengt.furberg@artimplant.se)

Adam Taube, professor, statistiker, institutionen för informationsvetenskap (statistik), Uppsala universitet (adam.taube@dis.uu.se)

Thomas Högborg, docent, överläkare, kliniken för gynekologisk onkologi, Universitetssjukhuset, Linköping (thomas.hogberg@ibk.liu.se)

Med P:n som i programpaketet

Om datatortyr och signifikansfiskeri

■ För studiet av prognostiska faktorer tycks en tradition eller, om man så vill, slentrian ha etablerats om hur sådana studier skall gå till.

Forskarna har undan för undan fått tillgång till avancerade statistiska metoder som till exempel multipel regressionsanalys, överlevnadsanalys (så kallade Kaplan–Meier-kurvor), logistisk regression, Cox-regression och »neurala nätverk« där olika bakgrundsvariablers inverkan på patienternas utveckling, till exempel recidiv eller överlevnad, kan studeras. Programmen ger dessvärre inga varningssignaler när materialen är för små, och givetvis slår de heller inte larm om själva grundmaterialet, till exempel de patologiska klassifikationerna, skulle vara felaktigt. Manualerna berättar om hur själva räknearbetet skall genomföras men utgör inte läroböcker i statistik.

I och med att statistiska programpaket blivit alltmer användarvänliga är det möjligt för forskare utan statistisk grundutbildning att välja och vraka bland de statistiska metoderna, precis som en sjuk person kan välja och vraka i sitt medicinskåp bland de piller och pulver som råkar stå på hyllorna. I motsats till en del piller och pulver är dock alla programpaket receptfria.

Med tiden vänjer sig läsarna av vetenskapliga artiklar vid de analysförfaranden som där tillämpas, oavsett om dessa är de mest lämpliga eller inte. Ingen tycks reagera nämnvärt när avancerade statistiska procedurer missbrukas på alldeles för små statistiska material. En metodologisk tradition har utvecklats i vårt land där avhandlingsarbeten om prognostiska faktorer studeras på alldeles för små material och resultaten sällan konfirmeras i oberoende större studier. Denna tradition ger en ineffektiv forskning som förmodligen genererar flera disputerade doktorer än användbara prognostiska faktorer.

Ett standardförfarande

Ett vanligt förfarande tycks vara att retrospektivt från något patientregister identifiera en grupp patienter med en viss diagnos, notera alla möjliga bakgrunds faktorer av eventuell betydelse samt lägga till några nya faktorer av intresse, för att sedan se efter i vad mån dessa är relaterade till patienternas överlevnad, tid till recidiv eller något annat utfallsmått. Därvid görs först det som brukar kallas »en univariat analys«, i praktiken en kartläggning av en serie bivariata »samband«, för det mesta belysta med regressionsanalys, en för varje en-

SAMMANFATTAT

För studiet av så kallade prognostiska faktorer använder de medicinska forskarna en rad avancerade tekniker, numera tillgängliga i statistiska programpaketet.

En tradition har utvecklats där överlevnad eller tid till återfall analyseras på grundval av data från statistiska material med få patienter men många möjliga förklarande variabler. Man gör »fishing expeditions« där p-värden används som betygsättning för potentiella förklarande variabler.

Eftersom materialen är små utfaller inte alla relevanta prognostiska variabler som signifikanta. Genom att många olika variabler provas blir risken stor för slumpsignifikanser.

Det räcker inte att visa att prognostiska faktorer är effektiva i den grupp där man först studerat dem, utan resultatet måste sedan verifieras på grundval av nya oberoende patientgrupper.

Evidensbaserad medicin

skild bakgrundsfaktor. Denna används därvid som förklarande variabel och överlevnaden eller tid till återfall som beroende. Varje analys, det vill säga varje möjlig förklarande variabel, »betygsätts« med ett p-värde. Om detta är litet antas variabeln kunna ha prognostisk betydelse.

Nästa steg är en multivariat analys där flera förklarande variabler får vara med på en gång. De får återigen varsitt p-värde. För det mesta visar det sig då att några av de variabler som tidigare hade statistiskt signifikanta p-värden nu får större icke signifikanta p-värden, det vill säga en potentiell prognostisk faktor tappar betydelse när samtidigt andra förklarande variabler beaktas eftersom dessa kan vara sinsemellan korrelerade. Detta är på intet sätt förvånande. Det finns sedan speciella metoder (stegvis regression m m) för att vaska fram

ett mindre antal förklarande variabler som kan påstås vara de viktigaste »oberoende« prognostiska faktorerna.

Trots den mängd p-värden som denna procedur genererar är det inte längre fråga om signifikantestning av konventionellt slag utan i stället »statistisk modellering«, en procedur som går ut på att vaska fram en matematisk modell, för det mesta i form av ett regressionsuttryck, som så bra som möjligt passar just till föreliggande data. Det är alltså fråga om statistiska modeller utan ytterligare förankring i biologisk teori än att de studerade variablerna – förhoppningsvis – är biologiskt meningsfulla. I fokus för intresset borde då stå mått på modellens totala anpassning. Om det är fråga om till exempel logistisk regression, där den beroende variabeln är dikotom (till exempel återfall/inte återfall eller död/inte död), eller Cox-regression finns det test och lämpliga mått på anpassning grundade på »likelihood«-kvoter att tillgå. En annan naturlig ansats kan vara att bilda mått på hur väl den anpassade regressionsekvationen predikterar utfallet för de olika individerna, jämfört med vad som verkligen inträffade. I fallet med logistisk regression blir grundvalen då en fyrfältstabla med det predikterade utfallet på ena ledden och det faktiska på en andra. Trots att det primära målet är att få fram en hel modell och alltså inte att göra signifikansbedömningar av vissa i förväg specificerade hypoteser, presenteras emellertid resultaten, enligt medicinsk statistisk praxis, med p-värden överallt där »samband« förekommer – ett i medicinsk litteratur utbrett missbruk av p-värden som påtalats av flera författare [1-3].

Det är inte ovanligt att en viss patientgrupp på detta sätt blir föremål för ett akademiskt doktorsarbete där man i en serie

uppsatser undersöker dels betydelsen av tidigare accepterade prognostiska faktorer, dels den eventuella nyttan av nya variabler, som för närvarande råkar vara »en vogue«. När materialet är för litet ursäktas detta med att det inte fanns flera patienter i det tillgängliga registret – med rimliga resurser gick det helt enkelt inte att skrapa ihop ett större material. Så kan det givetvis vara, men matematikens lagar tar dessvärre inte hänsyn till sådana omständigheter, och förfarandet leder till vad som i litteraturen betecknats som »data torturing« [4]. När väl avhandlingens alla uppsatser är klara vill ingen gärna bråka med doktoranden om att materialet var för litet, ingen vill försvåra för en person som följt handledarens instruktioner och gjort det bästa av situationen.

Många variabler och få observationer

För att belysa vissa principer och samtidigt undvika diskussion om medicinska detaljer utgår vi först från ett fiktivt, icke medicinskt exempel. Anta att en pedagog studerar data om 50 skolbarn i viss ålder. För var och en har noterats terminsbetyg eller någon annan mätare på framgången i studierna, vilken sedan skall »förklaras« av alla möjliga bakgrundsvariabler, här för enkelhets skull presenterade som alternativ. Det kan vara kön (pojke/flicka), kroppslängd (»långa«/»korta«), förekomst av syskon (enda barnet eller inte), föräldrarnas situation (ensamförälder eller par), nationalitet (svensk/invandrare), bostadsort (land/stad) etc.

Den »univariata analysen« innebär då helt enkelt att man först gör en jämförelse mellan betygen hos pojkar och flickor, sedan mellan »långa« och »korta« (kanske mått på mogenaden?), sedan mellan dem som är enda barn och dem som har syskon etc. Görs tillräckligt många sådana uppdelningar av materialet dyker det förr eller senare upp någon som ger en viss skillnad. Den kanske då avspeglar något intressant »samband«, men risken för att materialet på grund av en slump befinnes vara snett i något avseende och därmed genererar en statistiskt signifikant skillnad ökar givetvis med antalet studerade variabler.

Det är sedan dags att göra en »multivariat analys«, det vill säga ta med flera förklarande variabler samtidigt. Då inkluderas antingen samtliga möjliga förklarande variabler eller också endast de som blev signifikanta vid den första genomgången. Det förefaller rimligt att i vårt exempel endast beakta de uppdelningar som gav signifikans vid den »univariata« analysen. Anta att det blev signifikanta betygsskillnader vid uppdelningarna med avseende på kön, kroppslängd, nationalitet och bostadsort. Materialet sorterar därför med avseende på dessa fyra variabler samtidigt, vilket leder till en uppdelning i $2 \times 2 \times 2 \times 2 = 16$ klasser (»celler«). Eftersom materialet bara omfattade 50 individer kan det förväntas att några celler endast innehåller ett fåtal individer eller till och med är tomma. Att jämföra betygen mellan alla dessa 16 klasser blir alltså inte så enkelt.

Situationen underlättas inte särskilt mycket om man trycker på en viss teoretisk modell på materialet, det vill säga antar att det skall föreligga ett i förväg specificerat (matematiskt) mönster, vilket är en förutsättning för regressionsanalysen. Några enskilda celler är fortfarande glest befolkade. Det statistiska materialet är och förblir helt enkelt för magert för att tillåta slutsatser om alltför många variabler. Inte ens de mest sofistikerade statistiska tekniker kan kompensera för detta.

I ovanstående exempel hade man åtminstone ett värde – betyg eller vad det nu kan ha varit – för varje individ att »förklara«. När det gäller medicinska prognosdata studeras ofta tid till en viss händelse (»time to event«-data), till exempel återfall eller död i speciell sjukdom. Om det gäller till exempel data om 100 patienter, och ingen har hunnit få återfall eller avlidit, finns inget att »förklara«. Idealet från forsknings-

II Fakta 1

Checklista för studium av uppsatser om prediktiva faktorer

Från statistisk synpunkt bör läsaren vara observant på följande:

- Hur många patienter ingår i studiegruppen?
- Hur många »händelser« studeras?
- Hur många potentiella förklarande variabler studeras?
- Är det rimliga proportioner mellan antalet patienter/händelser och antal variabler?
- Justeras signifikansnivån med hänsyn till antalet genomförda test?
- Prövas de framtagna prediktiva faktorerna på nya, oberoende data?

Mer detaljerade checklistor återfinns i referens [12].

synpunkt är givetvis att patienterna har kunnat följas så länge att samtliga individer har avlidit (eller fått återfall), det vill säga man har uppnått maximalt antal händelser för den studerade populationen. Den information som materialet kan ge beror alltså inte bara på antalet studerade individer utan även på antalet händelser som inträffat, vilket i sin tur beror på uppföljningstiden.

Överväganden om materialstorlek

Rent intuitivt inses att det bör råda ett rimligt förhållande mellan materialets storlek och antal förklarande variabler som studeras, och olika tumregler har presenterats i metodlitteraturen. Sådana regler kan givetvis motiveras med teoretiska argument eller med hjälp av datorsimuleringar. En rekommendation som återfinns i vissa läroböcker formuleras av Petrie och Sabin [5] gällande vanlig linjär multipel regressionsanalys så att: »A usual rule-of-thumb is to ensure that there are at least 10 times as many individuals as explanatory variables.«

Ifråga om studier där tid till viss händelse står i fokus skärps detta krav ytterligare, så att enligt [6]: »The ratio of 10:1 relates number of events to number of prognostic variables studied; it does not represent the relationship between number of patients and number of variables finally selected.«

I samband med till exempel en klinisk prövning är en av hörnstenarna beräkningen av det antal observationer som fordras för att projektet skall få tillräcklig styrka (»power«) så att meningsfulla samband kan klarläggas med viss säkerhet. Denna typ av beräkningar tycks helt lysa med sin frånvaro i studier om prognostiska faktorer inom vissa medicinska områden, trots att metodik och rekommendationer står att finna i litteraturen, se till exempel [6].

Förutom överväganden om statistisk styrka finns en rad andra metodologiska aspekter [7], vilka ofta negligeras. En må dock speciellt nämnas – den vanliga underlåtenheten att justera för multipla jämförelser. Det finns, inte minst i litteraturen om kliniska prövningar, mycket skrivet om hur multipliciteten skall bemästras, till exempel genom justering av signifikansnivån med hjälp Bonferonis metod. I arbeten om prediktiva faktorer inom till exempel onkologin är det vanligt att helt bortse från denna problematik.

Några onkologiska exempel

Bland onkologiska arbeten där man söker efter bakomliggande prognostiska faktorer förekommer inte sällan uppsatser

med häpnadsväckande små patientmaterial som underlag för de statistiska analyserna. En härom året framlagd avhandling inleds med en uppsats, accepterad i välnummerad tidskrift, där ett material omfattande 49 patienter (35 händelser) först utnyttjas i elva »univariata« analyser och sedan för en multivariat analys med sex variabler. En undergrupp om 25 patienter specialstuderas ytterligare, och i uppsatsen redovisas allt som allt 27 p-värden. Det är inte förvånansvärt att författaren finner statistiska signifikanser. I samma avhandling finns ytterligare en uppsats, även den accepterad i välnummerad tidskrift, där data om 26 patienter får ligga till grund för 21 signifikansanalyser (sex jämförelser mellan två undergrupper och 15 så kallade Cox-analyser med en förklarande variabel i taget). Även här erhålls naturligtvis några statistiskt signifikanta samband, vilka inte nödvändigtvis behöver vara »sanna«. (Någon referens till den aktuella avhandlingen ges inte här eftersom syftet inte är att kritisera just det arbetet.)

Det bör tilläggas att denna avhandling på intet sätt är unik, utan den avspeglar ett arbetssätt som vunnit acceptans inom medicinska kretsar – det är lätt att hitta ytterligare exempel av samma slag. Det är förvånansvärt att ingen reagerar, men måhända har det att göra med att de statistiska signifikansanalyserna ger ett så vetenskapligt intryck att den genomsnittliga medicinska läsaren sväljer alltihop och – framför allt – har vant sig vid att så här går det till.

I vissa avhandlingsarbeten, där man med till exempel Cox-analys söker vaska fram alla »oberoende« prognostiska faktorer, delas arbetet upp i flera uppsatser (i olika tidskrifter) där olika grupper av förklarande variabler studeras på grundval av ett och samma datamaterial. I en nyligen ventilerad avhandling låg data om något över hundra patienter med drygt trettio händelser till grund för de fyra första uppsatserna (vilka innehöll sammanlagt över 80 p-värden).

Vad blir följden?

För det första bör än en gång framhållas att ju flera variabler som studeras, desto större är sannolikheten för att en statistisk signifikans skall dyka upp av en ren slump. När man systematiskt letar i ett speciellt datamaterial efter statistiskt signifikanta samband, det vill säga gör det som i litteraturen ibland kallas för »a fishing expedition«, får de i medicinska arbeten så populära p-värdena inte alls samma tyngd som i andra sammanhang eftersom signifikansbedömningen redovisas just därför att man råkat stöta på ett »statistiskt signifikant« p-värde och inte för att en speciellt i förväg uppsatt hypotes står i centrum för intresset.

För det andra har (för) små studier inte tillräcklig statistisk styrka för att eventuellt användbara prognostiska variabler – inklusive dem som redan är väl etablerade – verkligen skall slå igenom. En konsekvens av detta är att variabler som inte blir signifikanta i en sådan analys trots detta mycket väl kan ha en prognostisk betydelse. Påståenden om att vissa variabler saknar prognostiskt värde eftersom de inte fick statistisk signifikans kan därför vara direkt felaktiga.

För det tredje ger Cox-analyser eller logistiska regressioner baserade på för tok för små material en »vetenskaplig fasad« där läsaren inte får någon uppfattning om hur magert underlaget verkligen är.

Det har genom åren publicerats hyllmetrar om prognostiska faktorer inom onkologin. Trots detta är den kliniska användningen av prognostiska/prediktiva faktorer mycket begränsad. I ASCO (American Society of Clinical Oncology) antog man nyligen riktlinjer avseende användning av tumörmarkörer vid bröst- och koloncancer [8, 9]. Rekommendationerna var huvudsakligen negativa eftersom kvaliteten på det publicerade underlaget ofta inte var tillräcklig för slutsatser till grund för klinisk tillämpning. Detta kan sägas vara gans-

ka förvånande med tanke på att ASCOs riktlinjer huvudsakligen behandlar markörer som identifierades för många år sedan. På grund av detta publicerade ett antal av deltagarna i ASCO-kommittén en artikel där de beskrev ett graderings-system för tumörmarkörer [10] med metodologiska anvisningar för studier, som utvecklats ytterligare i en senare artikel [11].

Gion och medarbetare gav 1999 ut en guide för granskare (och författare) i the International Journal of Biological Markers [12]. I en anonym ledarkommentar två år senare i samma tidning [13] påpekas att trots dessa metodologiska rekommendationer är den statistiska kvaliteten fortfarande låg för artiklar publicerade i denna och andra medicinska tidskrifter med höga »impact factors«. En fråga inställer sig osökt: Om författarna lägger så liten vikt vid en så viktig del av en artikel som statistiken, hur är det då med övrig metodologi i de aktuella arbetena?

Hur skall man göra?

Det är givetvis fullt legitimt att detaljstudera en viss grupp av patienter för att söka få fram prognostiska faktorer. Samtidigt måste accepteras att ett litet material helt enkelt inte innehåller tillräckligt med information för ett studium av en stor mängd förklarande variabler. Detta beklagliga faktum bör inte skylas över med hjälp av sofistikerade statistiska analyser som läsarna inte kan genomskåda.

Om till exempel tre möjliga prognostiska faktorer skall studeras kan det vara klokt att börja med en något mindre prentiös ansats, till exempel att sortera upp materialet i undergrupper med hänsyn till alla dessa tre på en gång och se efter hur många individer som verkligen finns i de olika cellerna. Då blir det genast klart hur magert materialet kanske är. För varje cell med ett tillräckligt antal individer och händelser kan man sedan till exempel rita en överlevnadskurva (med konfidenstränser, se [14]!) och, som Gunnar Emanuel Sträng en gång uttryckte det, »ta sig en reflexionsfundering«. Om man sedan på grundval av data om en viss patientgrupp tror sig – på ett eller annat sätt – ha fått fram »optimala« prognostiska faktorer bör givetvis materialet även sorteras upp med avseende på dessa, och utfallet studeras i varje undergrupp.

Det kan givetvis hända att en forskare på grundval av ett visst patientmaterial lyckas, med ett intrikat statistiskt snickeri, få fram en statistisk modell som på ett helt enastående sätt passar till givna data. I så fall är det givetvis viktigt att konstatera att den verkligen gör det. Samtidigt bör med kraft understrykas att detta endast är ett första steg på vägen, forskaren har knappt nått halvvägs. Eftersom det aktuella patientmaterialet legat till grund för modellens härledning är det något av en självklarhet att anpassningen blir god. Det är först efter ett studium av hur den presenterade modellen fungerar för andra, oberoende patientgrupper, som metoden kan betygsättas på ett meningsfullt sätt. Tyvärr saknas detta sista fundamentala forskningssteg i en rad arbeten om prediktiva faktorer inom onkologin.

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Salsburg DS. The religion of statistics as practised in medical journals. *The American Statistician* 1985;39:220-3.
2. The end of P-values [editorial]? *British Heart Journal* 1988;60:177-80.
3. Taube, A. Varför alltid sätta P för signifikanta resultat? *Läkartidningen* 1990;87:960-2.
4. North RL. Data torturing. *N Engl J Med* 1993;329:1196-9.
5. Petrie A, Sabin C. *Statistics at a glance*. London: Blackwell Science; 2000.
6. Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. *Br J Cancer* 1994;69:979-85.
7. Levine MN, Browman GP, Gent R, Roberts R, Goodyear M. When is a prognostic factor useful?: A guide for the perplexed. *J Clin Oncol* 1991;9:348-56.
8. Clinical practice guidelines for the use of tumor markers in breast and colorectal cancer. Adopted on May 7 by the American Society of Clinical Oncology. *J Clin Oncol* 1996;14:2843-77.
9. Bast RC Jr, Ravdin P, Hayes DF, Bates S, Fritsche H Jr, Jessup JM, et al. 2000 update of recommendations for the use of tumor markers in breast and colorectal cancer: clinical practice guidelines for the American Society of Clinical Oncology. *J Clin Oncol* 2002;19:1865-78.
10. Hayes DF, Bast RC, Desch CE, Fritsche H Jr, Kemeny NE, Jessup JM, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* 1996;88:1456-6.
11. Hayes DF, Trock B, Harris AL. Assessing the clinical impact of prognostic factors: when is »statistically significant« clinically useful? *Breast Cancer Res Treat* 1998;52:305-19.
12. Gion M, Boracchi P, Biganzoli E, Daidone MG. A guide for reviewing submitted manuscripts (and indications for the design of translational research studies on biomarkers). *Int J Biol Markers* 1999;14:123-33.
13. Tutorial biometry corner [editorial]. *Int J Biol Markers* 2001;16:198.
14. Altman DG, Machin D, Bryant TN, Gardner MJ. *Statistics with confidence*. 2nd ed. Bristol: BMJ Books; 2000.

SUMMARY

”P” as in program package
Warning for spurious significances in studies of prognostic factors

Adam Taube, Thomas Högberg
Läkartidningen 2002;98:3302-5

For the study of prognostic factors, the medical researchers have access to a number of advanced statistical techniques available in standard program packages. A tradition has developed where survival or time to relapse is analysed on the basis of statistical materials with few patients but a large number of possible explanatory variables. In statistical »fishing expeditions« the p-values are used to sort out potentially useful prognostic variables. Since the number of observations is small, all relevant prognostic factors do not give statistical significance. Since a large number of variables are tested there is a considerable risk for spurious significances. It is not enough to show that a prognostic factor seems to be efficient in the patient group where it was first found. The result must be verified in further studies of independent groups of similar patients.

Correspondence: Adam Taube, Information Science, Uppsala universitet, Box 513, SE-751 20 Uppsala (adam.taube@dis.uu.se).

**D E B A T T
O M
E B M**

ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Gunnar Steineck, docent, överläkare, Klinisk cancerepidemiologi, institutionen för onkologi och patologi, Karolinska institutet och Samhällsmedicin, Stockholms läns landsting, Stockholm

Jan Adolfsson, docent, överläkare, KARO-institutionen, Karolinska institutet och Statens beredning för medicinsk utvärdering (SBU), Stockholm

Fixeringen vid randomisering hämmar klinisk behandlingsforskning

Att det randomiserade kliniska försöket medverkat till utvecklingen av en vetenskapligt baserad sjukvård råder det inga tvivel om, men de två artiklar om »randomiseringens välsignelse« som publicerats i Läkartidningen [1, 2] kan försvåra arbetet med att anta dagens stora utmaningar för den kliniska forskningen. Att sätta randomiseringen som en kvalitetsstämpel, som kvalitativt skiljer ut sådana studier från andra vad gäller validitet, har som grund en förenklad syn på svårigheterna att nå valida resultat inom klinisk forskning. Det visar en nyligen gjord sammanställning, där resultaten från randomiserade och icke-randomiserade studier jämförs [3-5]. Att randomiseringen ger en studie specifika förutsättningar för statistisk inferens, som inte finns i icke-randomiserade studier, är matematiskt riktigt, men relevansen för praktisk klinisk forskning kan ifrågasättas. För att möta dagens kliniska problem behöver vi studieutformningar som är tillräckligt valida för att kunna urskilja betydelsefulla samband för kliniskt prioriterade frågeställningar. Sådana ansträngningar försvåras om en matematisk idealvärld ställs i kontrast till undersökningarnas resultat. Fruktbart är i stället att bedöma validiteten för varje enskild undersökning – utan författade meningar om att en viss utformning (t ex randomisering) medför kvalitet.

Långtidseffekter av kirurgi

Mortaliteten vid många stora kirurgiska ingrepp har minskat nästan till gränsen för vad som är möjligt att uppnå. Utvecklingsarbetet kan därför koncentreras på att undvika kroniska oönskade symtom som en följd av det kirurgiska traumat. Hur skall då en utvärdering av långtidseffekter efter kirurgi gå till? En randomisering av patienterna mellan operation och icke operation är den första tanken, men vad gäller tumörkirurgi är det två effekter som studeras vid en uppföljning efter tre år, den av kirurgin och den av den växande tumören. Här kan ingen självklar inferens till en behandlingseffekt göras. (Dessutom är studien idag olämplig och oetisk för många frågeställningar.) Ett alternativ vore att studera operationen med en före-efterdesign, men då har vi återigen problem med att en växande tumör, liksom åldrandet, stör inferensen. Då återstår att randomisera friska individer mellan tumörkirurgi och ingen tumörkirurgi. Vi har då närmat oss

SAMMANFATTAT

Fixering vid randomisering kan minska effektiviteten i klinisk behandlingsforskning.

En studie kan indelas i fyra steg, i varje steg kan förödande validitetsbrister uppkomma. Randomisering motverkar huvudsakligen brister i ett av stegen.

Icke-randomiserade studier kan vara konklusiva (fastställa kausalitet) och randomiserade studier icke-konklusiva.

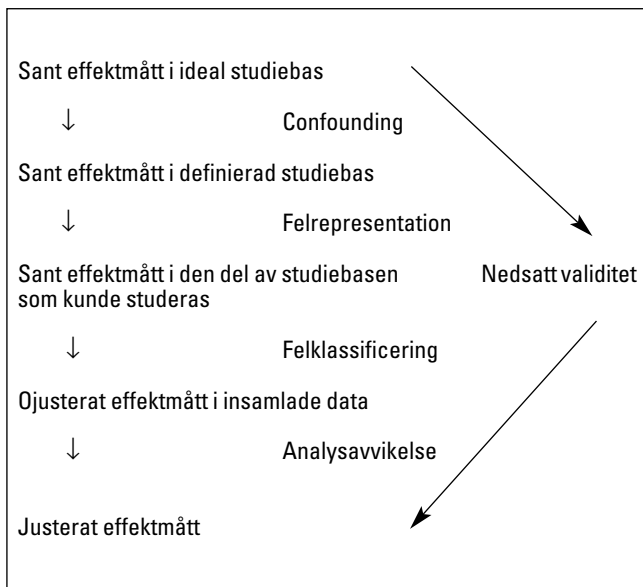
Randomisering, placebo-kontroll, samtidig start av utfallsobservation, frånvaro av bortfall samt trippelblindning (patient, behandlare och utvärderare) bör alltid eftersträvas eller efterliknas

Serie Evidensbaserad medicin

idealet – om enbart validiteten beaktas. Statistisk inferens är möjlig (om studien i övrigt är valid). När statistiskt signifikanta skillnader erhålls kan vi med säkerhet uttala oss om de kroniska effekterna av tumörkirurgin. Naturligtvis varken kan eller bör en sådan undersökning utföras. Lyckligtvis finns ett praktiskt möjligt alternativ: en jämförelse mellan opererade individer och befolkningskontroller. En sådan studie har lämnat den matematiska idealvärlden, men den kan ge medicinskt relevant kunskap och är för vissa problem den mest valida av de studieutformningar som kan genomföras.

Radikal hysterektomi

I en svensk studie identifierades samtliga fall av »tidig« livmoderhalscancer (Ib-IIa) som registrerades i Sverige 1991–1992 via landets sju gynekologiska centra [6]. Patienterna matchades (ålder, boende i samma kommun) mot kvin-



Figur 1. Fyra steg i en klinisk studie och den subkategori av systematiska fel som påverkar validiteten i varje steg. Modifierat efter Steineck och medarbetare 1998.

nor i den allmänna befolkningen. En relativ risk på 6,1 (95 procents konfidensintervall 2,6–14,1) noterades för vaginal korthet (relaterat till smärtsamma samlag) då resultaten för de tidigare patienterna (cirka fem år efter operationen) jämfördes med befolkningskontrollerna. Studien uppfyller inte alla kriterierna för statistisk inferens, ändå hävdar författarna att ett orsakssamband föreligger mellan operationen och det kroniska oönskade symtomet. Detta är provocerande för den som ser randomiseringen som ett sine qua non för inferens.

Fyra steg i studie

En klinisk undersökning kan ses som fyra steg från den ideala studien till det effektmått (relativ risk eller riskskillnad då två grupper jämförs) som beräknas [7]. För varje steg ger ett klassiskt kliniskt försök, dvs randomisering, placebo, inget bortfall, observationsstart i relation till randomiseringsdatum, trippelblindning (patient, behandlare, utvärderare) [8], uppslag till hur validiteten kan höjas, och valideringsteori [9] ger rättesnören för hur de systematiska fel som kan uppstå i steget påverkar effektmåttet (Figur 1). I det första steget, då studiebasen definieras, kan randomisering och placebo, eller matchning liksom andra kvasiexperimentella angreppssätt, höja validiteten. De systematiska fel som uppstår påverkar effektmåttet enligt reglerna för störfaktorer. I det andra steget, då delar av studiebasen förloras, undviks systematiska fel genom blindning och minimerat bortfall, och de lyder under reglerna för felrepresentation (bristande representation av studiebasen). Under det tredje steget, då data insamlas (exempelvis genom mätning av utfallet), kan blindning höja validiteten och systematiska fel påverka effektmåttet enligt reglerna för felklassificering. Slutligen bör analyser göras enligt »state of the art«, och även i detta steg kan validiteten minskas (analysavvikelse). Om inte maximal validitet kan uppnås i ett steg bör metoderna i det klassiska kliniska försöket efterliknas (Tabell 1).

Icke-randomiserad konklusiv studie

Att i den komplexitet som validitetsbedömningen av en klinisk studie innebär enbart betona randomiseringen i det första steget kan leda fel. Under varje steg finns möjligheter att införa förödande validitetsbrister. Men i varje studie, icke-ran-

domiserad som randomiserad, finns dessbättre i varje steg även möjligheter att tillförsäkra effektmåttet en tillräcklig validitet genom att på olika sätt kompensera för avstegen från metoderna inom ett klassiskt kliniskt försök. I exemplet ovan gällande radikal hysterektomi använde författarna under steg 1 populationsbaserade register för att undvika inbyggd selektionsproblematik i studiebasen. Vidare insamlades information om ett stort antal potentiella (direkta eller ställföreträdande) störfaktorer, och deras inverkan på den relativa risken undersöktes. För validiteten i steg 2 gjordes upprepade förstudier för att minimera bortfallet, och blindning simulerades. Trippelblindning efterliknades (ökar validiteten även i steg 3) genom att standardiserade frågor besvarades i hemmet utan påverkan av utvärderare (intervjuare), genom att instrumentet insändes anonymt (motverkar skeva förväntningar) och genom att individerna inte visste för vilka utfall en effekt väntades. Slutligen utfördes steg 4 med etablerade statistiska metoder. En identisk metodik kan användas för att utvärdera en omvårdnadshandling [10].

Randomiserad icke konklusiv studie

En förödande validitetsbrist kan fås i steg 1 om okända riskfaktorer för utfallet är ojämnt fördelade mellan de grupper som jämförs, om ytterligare interventioner stör den behandling som studeras eller om en utfallskurva utläses skevt [11]. Ett selektivt bortfall (risken ökar om det är stort) kan ge avgörande brister i steg 2 [9, 12, 13], en varierande diagnostisk intensitet (t ex genom att symtom dokumenteras av en utvärderare som känner till behandlingen) i steg 3 [8, 12, 14], och rena räknefel liksom felaktiga statistiska modelleringar i steg 4 [12]. För den ovan refererade frågeställningen huruvida radikal hysterektomi ger kroniska oönskade symtom är det lätt att konstruera en randomiserad studie med någon av dessa förödande brister (ingen randomiserad studie pågår i verkligheten); en sådan studie skulle således vara inkonklusiv. Att sätta randomiseringen som det främsta validitetskriteriet kan hindra de verkliga möjligheterna att nå kunskap.

Effektiviteten i klinisk behandlingsforskning måste öka om vi skall kunna möta det tryck från patienter som nu uppkommer via Internet och patientföreningar när arbetet med kvalitetsregister och vårdprogram visar en väsensskild praxis mellan olika centra. Den kliniska forskningen måste möta kraven på ett underlag för rationella val av terapi inom en rimlig tidsperiod. Det kan ske genom att icke-randomiserade studier utnyttjar den kvasiexperimentella situation som föreligger därför att olika regioner i landet behandlar patienterna helt olika. Effektiviteten måste öka också därför att patienterna får kännedom om, och producenterna önskar sälja, ett stort antal nya medel – vid det senaste årliga stora mötet inom cancerområdet (ASCO) räknades substanserna i hundratal. Denna nya situation fordrar troligen nya, innovativt (randomiserade eller icke-randomiserade) utformade Fas II-studier för att bli hanterbar. Vidare måste effektiviteten ökas för att ta fram valid kunskap för de interventioner som av etiska, praktiska (för lite tid, för få nya fall) eller ekonomiska skäl inte kan utvärderas med en konventionell randomiserad studie. Det gäller i första hand långtidseffekter av kirurgi, enskilda läkemedel eller kombinationsbehandlingar, omvårdnadshandlingar eller information, men även effekter på kort sikt av vissa interventioner. Behovet ökar även därför att patienter i minskad utsträckning vill bli lottade till sin behandling, speciellt om många av utfallen (t ex frekvensen av fem olika symtom) för behandlingsalternativen till viss del är kända. Icke-randomiserad utvärdering behövs då effektiviteten av information, omvårdnadshandlingar eller medicinska åtgärder såsom de utförs inom den faktiska sjukvården, utanför definierade behandlingsprotokoll, skall undersökas. Slutligen ställer finan-

Tabell I. Aspekter på validitet före, under och efter en klinisk studie. Modifierad efter Steineck och medarbetare, 1998.

Utformning som kan öka validitet	Genomförande som kan öka validitet	Bedömning av validitet då studien genomförts
<i>Confounding:</i> Randomisering (slumpmässig fördelning av behandling till olika individer)	<i>Confounding:</i> Försäkran om att slumpmässighet föreligger	<i>Confounding:</i> Undersökning om slumpmässighet förelåg
»Före–efterdesign« – mellan grupper eller individer Matchning (behandlad– obehandlad) för störfaktorer – på individ- eller gruppnivå Utnyttjande av en kvasiexperimentell situation	Försäkran om gruppstillhörighet	Undersökning i vad mån information insamlats om kända möjliga störfaktorer Bedömning av storleken på kvarvarande confounding efter justering (via »change-in-estimate« och bedömning av stör-faktorns mätfel) Bedömning om kända eller okända möjliga störfaktorer kan föröda validiteten. Undersökning om studiebasen har inbyggd selektionsproblematik (att utfallet påverkat definitionen)
Blindning av patient och vårdgivare (placebo) Shambehandling – gör behandling lika i de grupper som jämförs (utom den behandling som studeras)	Försäkran om att individerna inte vet vilken av de alternativa behandlingarna som givits	Bedömning om placebo och shambehandling motsvarar alla faktorer som kan påverka utfallet, utom den behandling som studeras (kirurgi, läkemedel, omvårdnadshandling)
Start av observationsperiod som bestäms av randomiseringstillfället Start av observationsperiod som relateras till insjuknande (eller dagen då ett visst sjukdomsstadium uppnåtts)	Noterande av randomiseringsdagen Beräkning av dag för insjuknande (eller dag då visst sjukdomsstadium uppnåtts)	Undersökning om avläsningen förskjutits i ena gruppen, »lead time«
<i>Felrepresentering:</i> Blindning av patient, vårdgivare och utvärderare (trippelblindning) för att undvika selektivitet i bortfallet Efterlikning av blindning Användande av metoder som minskar bortfall (inklusive exkludering)	<i>Felrepresentering:</i> Undvikande av bortfall (av individer eller observationstid (persontid) för viss individ)	<i>Felrepresentering:</i> Bedömning av storlek på bortfall. Bedömning om bortfallet är selektivt (att sambandet mellan behandling eller annan exponering) och utfall varierar mellan den del av studiebasen som är kvar och den som fallit bort
Vid fall–kontrollbeskrivning av studiebasen: val av kontroller som avspeglar fördelningen av den faktor (t ex behandling eller »exponering« som tumörkaraktistika) som studeras	Undvikande av bortfall	Som ovan
Vid sammanfogning av delstudiebas till en stor, »metaanalys av primär-material«: ambition att medtaga alla (oberoende av publicering)	Undvikande av bortfall av delstudiebas	Som ovan
Vid sammanställning av informationen i publicerade studiebas, »meta-analys av publicerade studier«: ambition att medtaga alla	Undvikande av bortfall av publicerade data	Som ovan
<i>Felklassificering:</i> Blindning av utvärderare för att undvika selektivitet i mätfel Efterlikning av blindning Användande av metoder för mätning av effekter (eller exponering) som minskar mätfelens storlek	<i>Felklassificering:</i> Försäkran om att blindning föreligger Undvikande av mätfel	<i>Felklassificering:</i> Bedömning av storlek på mätfel av utfallet och om det varierar mellan de grupper som studeras Bedömning av storlek på mätfel av exponering (om aktuellt i studien) och om det varierar med utfallet
<i>Analysavvikelse:</i> Ambition att insamla information om samtliga störfaktorer (speciellt vid icke-randomiserade studier). Ambition att insamla information om samtliga variabler som avbildar fördelningen av bortfall eller mätfel. Ambition att mäta störfaktorer korrekt. Ambition att använda statistiska modeller med hög »goodness-of-fit«	<i>Analysavvikelse:</i> Undvikande av rena räknefel. Enbart inkludering i statistiska modeller, »justering«, av variabler som ökar validiteten	<i>Analysavvikelse:</i> Bedömning om rena räknefel gjorts. Bedömning om effektmåttet enbart justerats för relevanta variabler Bedömning av »goodness-of-fit«

siärerna av sjukvården allt större krav på en dokumentation av produktivitet samt medicinsk och ekonomisk effektivitet i verksamheten – en kunskapsproduktion som har icke-randomiseringen inbyggd.

Utveckling i rätt riktning

En snabb utveckling inom klinisk forskningsteori, jämte empiriska studier som belyser validiteten i genomförda undersökningar, kan gynna icke-randomiserad klinisk behandlingsforskning. Arbetet pågår [3, 4], och de slutsatser som nås är kontroversiella [5]. Men även om vi kan hämta inspiration utanför Norden har vi egna förutsättningar för ett nytänkande med populationsbaserade register, många fall av »geografisk pseudorandomisering« (varierande behandlingspolicy i olika upptagningsområden) samt en välutbildad och samarbetsvillig befolkning som kan ge medicinsk information via enkäter. Dessa förutsättningar ger oss möjligheter till internationellt konkurrenskraftiga bidrag till utvecklingen.

Den senaste generationens preparat vid HIV-infektion godkändes som läkemedel utan att fabrikanterna hade genomfört randomiserade studier som visat positiva effekter på patienters överlevnad. Myndigheterna blev tvungna att prioritera tidsfaktorn, och exemplet visar att effektivitetskraven för klinisk forskning redan är tvingande. I bristen på randomiserade studier tvingas vi nu studera preparatets positiva och negativa kliniska effekter i den icke-randomiserade praxis som uppkommer, att argumentera för randomiserade studier skulle enbart försena utvecklingen.

Kalla realiteter

Att ifrågasätta randomiseringen som det främsta validitetskriteriet är långt ifrån en fråga om en esoterisk akademisk diskurs. Den som försöker driva icke-randomiserad behandlingsforskning möts snart av finansiärer eller tidskriftsredaktioner som kontrasterar den planerade eller genomförda studien mot en matematisk idealvärld. Redan det faktum som Adam Taube antyder [2], att det verkliga livets studier sällan eller aldrig ger matematikens förutsättningar för statistisk inferens, gör att randomiseringsfixeringen borde förpassas till 1900-talets konceptkyrkogård när det gäller praktisk klinisk forskning. Att hålla fast vid uppfattningen skulle ge den kliniska behandlingsforskningen en järnklump runt foten som försvårar de språng som måste tas mot framtiden. En rad frågeställningar kräver studier som är randomiserade, innefattar placebo för att uppnå dubbel- eller trippelblindning, är utan selektiva exklusioner och med lågt bortfall. När de är praktiskt möjliga att genomföra är de klassiska kliniska försöken att föredra, och de bör alltid efterliknas. Men det är den valida icke-randomiserade jämförande studien som kan välsigna 2000-talets kliniska behandlingsforskning.

Referenser

1. Nilsson G, Sörensen S. Leve den randomiserade studien! *Läkartidningen* 2000; 97: 3445-9.
2. Taube, A. Om randomiseringens välsignelser. *Läkartidningen* 2000; 97: 4173-4.
3. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *New Engl J Med* 2000; 342: 1878-86.
4. Concato J, Shan N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *New Engl J Med* 2000; 342: 1887-92.
5. Pocock SJ, Elbourne DR. Randomized trials or observational tribulations? *New Engl J Med* 2000; 342: 19007-9.
6. Bergmark K, Åvall-Lundqvist E, Dickman P, Henningsohn L, Steineck G. Vaginal changes and sexuality in women with a history of cervical cancer. *N Engl J Med* 1999; 340: 1383-9.
7. Steineck G, Kass PH, Ahlbom A. A Comprehensive clinical epidemiological theory based on the concept of the source person-time and four distinct study stages. *Acta Oncol* 1998; 37: 15-23.
8. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I et al. Improving the quality of reporting of randomized controlled trials. *JAMA* 1996; 276: 637-9.
9. Norell S. A textbook of epidemiology. New York: John Wiley and Sons, 1995.
10. Rådestad I, Steineck G, Nordin C, Sjögren B. Psychological complications following stillbirth – the influence of memories and immediate management: Population Based Study. *BMJ* 1996; 312: 1505-8.
11. Steineck G, Adolfsson J. Bias due to observation of different parts of a nonconstant hazard curve over time. *Urology* 1996; 48: 105-9.
12. Rothman KJ, Greenland S. Modern epidemiology. Washington: Lippincott and Raven, 1998.
13. Rådestad I, Olausson POP, Steineck G. Measuring errors and non-participation in a nationwide study of stillbirth. *Acta Obstet Gynecol Scand* 1999; 78: 592-8.
14. Steineck G, Adolfsson J, Whitmore WF Jr. »Local recurrence« and »disease-free survival«; doubtful parameters when comparing non-randomized studies of prostate cancer. *Scand J Urol Nephrol* 1991; 138: 121-6.

Kommentar 1

Varning för övertro på vetenskap

II Vår tid präglas av ett mycket stort förtroende för s k evidensbaserad medicin. Enligt evidensbaserad medicin ger resultat av randomiserade studier de säkraste riktlinjerna för medicinskt handlande medan resultat av s k observationsstudier (kohortstudier och fall-kontrollstudier) ger mindre tillförlitlig information. Steineck menar i artikeln ovan att respekten för randomiserade studier blivit för stor och utgör ett hinder för en sund medicinsk utveckling. Han menar att t ex jämförelse mellan besvär hos matchade kontrollpersoner i en allmän befolkning och besvär i en grupp patienter som hysterektomerats för tidig livmoderhalscancer ger lika säkra resultat som randomiserade studier.

Observationsstudier bekräftade i randomiserade studier

Emellertid har alla observationsstudier en grundläggande svaghet: designen är inte experimentell. Den enskilde patientens behandling är inte slumpvis vald som vid randomiserade studier, och följaktligen finns en oundviklig risk för selektionsbias och systematisk, av behandlingen oberoende, skillnad i behandlingsresultatet. Visserligen kan man i dataanalysen justera för kända faktorer av betydelse för behandlingsresultatet, men man kan inte justera för de många okända faktorer som påverkar detta resultat. Vi har goda exempel på värdefulla observationsstudier, vars resultat bekräftats i randomiserade studier och därmed drivit medicinsk utveckling framåt. Nämnas kan observationserna av ett starkt negativt samband mellan acetylsalicylsyra och hjärtinfarkt samt av ett negativt samband mellan folsyramedicinering i graviditetens första trimester och ryggmärgsbräck.

Å andra sidan finns flera exempel på att observationsstudier givit missledande fynd som kunnat avfärdas efter randomiserade studier. Sådana exempel är fynd av att vitamin E skulle förhindra kardiovaskulär sjukdom och att betakaroten skulle förhindra lungcancer. Det kanske mest spektakulära exemplet är HERS-studien (heart and estrogen/progestin replacement study) [1]. I denna studie kunde man avfärda tidigare iakttagelser från observationsstudier av god sekundärpreventiv effekt av östrogen vid kranskärlssjukdom. Den san-

nolika förklaringen är att östrogenkonsumerande kvinnor skiljer sig från andra kvinnor beträffande skyddsfaktorer mot kranskärslsjukdom, t ex kost- och motionsvanor och att detta förhållande misslett i observationsstudierna. Den randomiserade studien hindrar i detta fall en mycket omfattande och kostsam överbehandling med östrogen.

Förnuftig rimlighetsbedömning på kunskapsbas

Vi bör alltså vara tacksamma för den randomiserade studien. Å andra sidan bör vi vara ödmjuka när det gäller all kunskap och inse att den även efter goda randomiserade studier kan vara ett styckverk. Vi får akta oss för vidskeplig övertro på vetenskap. Den randomiserade studiens patienter är t ex ofta inte alls representativa för den patient man har mitt emot sig i undersökningsrummet.

Bristande förståelse för patofysiologiska mekanismer gör att vi lätt klumpar ihop heterogena patientgrupper på ett missledande sätt. Om man använt B₁₂ i en randomiserad studie på patienter med anemi skulle alla anemipatienter fått denna medicin om vi inte haft kunskap i anemiens patofysiologi.

Ibland är det praktiskt omöjligt att genomföra randomiserade studier, och observationsstudier är då av värde som ett näst bästa alternativ. Många kliniska situationer belyses inte alls av randomiserade studier eller observationsstudier och fordrar en förnuftig rimlighetsbedömning på basen av en relativt torftig tillgänglig kunskapsbas. Det är läkarens plikt och ansvar att ge även dessa patienter en bra vård.

Göran Nilsson

docent, överläkare, Hjärtsektionen, Centrallasarettet, Västerås

Referens

1. Hulley S, Grady D, Bush T, et al. Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women. JAMA 1998; 280: 605-13.

Kommentar 2

Randomiserade studien fortsatt gyllene standard

II Gunnar Steineck har rätt i att det ibland finns en fixering vid den randomiserade kliniska studien som den vetenskapliga metoden. En sådan fixering kan leda till en ensidig syn på klinisk forskning och dess uppgifter. Jag förstår emellertid inte helt hur hans metoddiskussion i övrigt relaterar till denna kritik. Randomiseringen har verkligen många välsignelser, och den randomiserade kliniska prövningen måste för närvarande anses vara vår gyllene standard för att utvärdera olika interventioner.

Randomisering i sig inget kvalitetsmått

Steineck framhåller helt riktigt att en studies kvalitet inte enkelt kan avgöras genom att konstatera om den är randomiserad eller inte. En klinisk studie kan vara mycket innovativ och metodologiskt avancerad och det bästa man kan göra under rådande omständigheter – alltså ha en mycket hög kvalitet – utan att vara randomiserad. En randomiserad studie kan dock vara undermåligt utförd och ge opålitliga resultat. Precis som Steineck framhåller måste man i stället bedöma den vetenskapliga kvaliteten och validiteten i varje enskild undersök-

ning för sig – utan förutfattade meningar om att en viss utformning medför kvalitet. Det är således en förenkling att betrakta värdet av studier utifrån en sorts vetenskaplig bevisvärdeshierarki som baseras kort och gott på studiedesign. Den av Steineck citerade debatten i New England Journal of Medicines utgåva från 22 juni 2000 (nummer 25, volym 342) demonstrerar detta. Båda inläggen är mycket läsvärda och har sina poänger.

Randomisering inte alltid möjlig eller nödvändig

I många viktiga kliniska studier kan man inte göra en randomisering. Detta gäller t ex studier av olika prognosfaktorer. Om man till exempel vill studera hur ålder vid diagnos påverkar prognosen vid bröstcancer, kan ju inte patienter randomiseras till att få sin diagnos vid olika ålder. Den moderna epidemiologin har ägnat decennier åt utveckling av design och analysmetoder för att göra så valida studier som möjligt i situationer där man inte kan göra ett kontrollerat experiment. Steineck har helt rätt i att denna metodologi bör tas till vara bättre. Det skulle vara mycket fruktbart om man överflyttade metodologiska landvinningar från etiologisk forskning till den kliniska epidemiologins domäner.

Andra typer av frågeställningar bjuder på etiskt mycket svåra problem vid randomisering. Ytterligare andra situationer gäller frågor där vi vill ha uppfattningar om interventioners effekter men inte anser att det är kostnadseffektivt med en omfattande och långvarig randomiserad studie. I ytterligare andra sammanhang finns redan erfarenheter i kliniska dataregister som måste tas till vara på ett effektivt sätt in- till dess att vi har resultat från planerade eller pågående randomiserade studier.

Steinecks val av exempel är emellertid i mina ögon inget som i grunden problematiserar användningen av randomiserade studier och skjuter därför litet bredvid målet. Just studier av oförutsedda biverkningar eller biverkningar på lång sikt som man inte övervägde vid behandlingstillfället lämpar sig nämligen extra väl för att angripa med icke-randomiserade studiedesigner [1]. Vid studier av oförutsedda biverkningar så har urvalet till behandlingsgrupp för det mesta skett utan just sikte på denna end-point. Urvalet har i stället oftast skett med hänsyn till förväntade positiva effekter eller med hänsyn till kända – och oftast kortsiktiga – biverkningar.

När det gäller bra studier i kliniska databaser har vi i Sverige som Steineck påpekar utomordentligt goda förutsättningar som vi bör utnyttja i mycket större utsträckning än vad som görs för närvarande.

Modetänkande alltid farligt

Konformistiskt tänkande och användning av olika argument som slagträn snarare än som rationellt avvägda skäl för eller emot är alltid skadligt. Karl Popper har i en kritik av just modetänkande uttryckt att människans historia kan ses som en rad katastrofala utbrott av modetänkande inom religionens, politikens och filosofins områden. Inom klinisk forskning har paradigmet randomiserad prövning ibland tenderat att bli ett sådant modetänkande. Att utan djupare analys förkasta studier därför att de inte är randomiserade gör att vi avhänder oss mycket stora mängder av värdefull information. Ibland förkastas till och med en randomiserad studie bara därför att den inte är 100-procentigt perfekt gjord. Sådana diskurser kan ju bli vanligare genom framväxten av en kader av metodologer med liten egen erfarenhet av klinisk forskning och som inte vet hur svårt eller omöjligt det är att göra den perfekta studien.

Länge leve den randomiserade studien!

Jag håller alltså med Steineck när han kritiserar ett modetänkande. Mina slutsatser blir dock inte riktigt desamma. Det

vore naturligtvis ett lika tokigt modetänkande – eller faktiskt värre – om man drastiskt skulle tona ner rollen av den randomiserade kliniska studien.

Jag tycker således inte att vi bör ifrågasätta lovsången om randomiseringens välsignelse i stort, men kritisera en överdriven fixering vid randomiseringen som metod. Jag tror inte att det är den valida icke-randomiserade jämförande studien som kan välsigna 2000-talets kliniska behandlingsforskning, utan den randomiserade kliniska studien kommer under över-skådlig tid att fortsätta att vara den gyllene standarden.

Steineck har därför helt rätt när han säger »när det är praktiskt möjligt är de klassiska kliniska försöken att föredra, och de bör alltid efterliknas«. Emellertid – där dessa försök inte är möjliga eller under väntan på deras resultat – så kan vi använda epidemiologins metodologiska framsteg för att på ett kraftfullt sätt befrukta den kliniska forskningen. I den meningen håller jag helt med Steineck. Det vore en välsignelse för klinisk forskning om alla de deskriptiva fallserier som idag fortfarande publiceras kunde ersättas av studier med en genomtänkt design och analys. Ofta utnyttjas tyvärr inte värdefulla data i kliniska material på ett effektivt sätt. Ibland ser man material som skulle kunna ta klivet från fallserie till mera analytiskt inriktad mycket informativ studie bara genom konsultation med personer erfarna i klinisk epidemiologi och genom anskaffande av en större studiebas genom att slå sig ihop med andra kliniker.

Lars Holmberg

docent, överläkare,

Regionalt onkologiskt centrum, Akademiska sjukhuset, Uppsala

Referens

1. Miettinen O. The need for randomization in the study of intended effects. *Stat Med* 1983; 2: 267-71.

Särtryck

Läkartidningen

När Försäkringsmedicinska Sällskapet bildades för att främja försäkringsmedicinens utveckling samlades 14 artiklar publicerade i Läkartidningen 1996 till ett särtryck. Detta belyser hur försäkringsläkare arbetar inom allmän och privat försäkring och tar upp försäkringsmedicinska problem från patientens synvinkel.

Riskbedömning vid barnförsäkringar, etiska problem i samband med gentestning och försäkring, samt de kniviga ärenden som gäller nack-skulderbesvär, inklusive pisksnärtskador, behandlas bland annat i artiklarna.

Priset är 50 kronor.

Försäkringsmedicin



Beställer härmed.....ex
av "Försäkringsmedicin"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till Läkartidningen
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck. böcker

Varning för alltför ensidig evidensbaserad medicin!

God läkarkonst förutsätter även människokänedom och intuition

II För åtskilliga år sedan behandlade Bertil von Ahn angina pectoris-patienter på Eksjö lasarett med ett s k rensningsdropp bestående av heparin med tillsats av nitroglycerin. Denna behandling ansågs lösa upp proppar i kranskärnen och vidga dem. Patienterna var i stort sett nöjda med behandlingen, som associerade till deras vardagserfarenhet av igenslamning av och flödeshinder i olika typer av rör. Behandlingen hade säkert en betydande placeboeffekt, men byggde också på patofysiologisk kunskap om heparinets koagulationshämmande effekt och nitroglycerinets kärlvidgande effekt. Båda dessa effekter var vid denna tid väl belagda. Några randomiserade studier av behandlingsutfallet fanns dock inte.

Kritiskt lagda nybörjare inom läkaryrket, även jag, var väl medvetna om denna brist och satte åtskilliga frågetecken för proceduren. 1990-talets randomiserade studier [1] har dokumenterat värdet av heparin vid instabil angina pectoris. I efterhand måste rensningsdroppen bedömas vara en bra terapi, även om vi numera kan göra denna behandling mer praktisk genom att heparin ges subkutant i lågmolekylär form.

På 1940-talet injicerade lasarettsläkaren Eskil Kyhlin i Jönköping hypofyseextrakt från nyss slaktade kalvar på patienter med förmodad hypofysinsufficiens. Dessa hade i många fall en klinisk bild av anorexia nervosa [Sven Nilsson, f d klinikchef, medicinkliniken, Jönköping, pers medd, 2000]. I cirka hälften av fallen förbättrades patienten med återupptagna menstruationer. Liksom implantationen av aptestiklar på parisiska dignitärer på 1920-talet tycktes denna behandling grunda sig på rimlig endokrinologisk experimentell kunskap. Behandlingsprinciper av detta slag tillämpas inte längre och har mig veterligen aldrig prövats i randomiserade studier.

Vad kan vi lära av medicinhistorien?

Dessa medicinhistoriska utflykter är intressanta med tanke på vår tids fokusering på s k evidensbaserad medicin, för vilken den randomiserade studien uppfattas som guldstandard för kliniskt handlande. Detta synsätt kan leda till en nedvärdering av medicinska åtgärder som är grundade på patofysiologisk kunskap, beskrivning av mindre patientserier och personlig erfarenhet. Läkarens handlande måste bygga på en syntes av alla dessa kunskapskällor. Att t ex trombolysbehandling ges vid akut hjärtinfarkt beror *både* på att trombolytiska medel lö-

SAMMANFATTAT

I begreppet evidensbaserad medicin spelar randomiserade studier med rätta en central roll. Randomiserade studier har dock svagheter.

Evidensbaserad medicin alltför ensidigt utgående från randomiserade studier riskerar att ge en instruktionsboksliknande medicinsk inriktning som kommer i konflikt med god läkarkonst.

Patofysiologisk kunskap, iakttagelser vid sjuksängen och personlig patienterfarenhet är oundgängliga förutsättningar för god läkarkonst.

Sinne för det speciella i varje klinisk situation, god allmän människokänedom och intuition fordras för att uppnå goda behandlingsresultat.

Evidensbaserad medicin

ser upp proppar i kranskärnen *och* på att diverse ISIS- och GUSTO-studier påvisar mindre mortalitet och morbiditet med dessa medel.

Enligt en substudie av ISIS-2 [2] har aspirin ingen effekt på hjärtinfarktpatienter födda i tvillingarnas och vågens stjärnbilder. Vi underlåter dock inte att ge aspirin till dessa patienter, eftersom vi helt enkelt inte kan förstå någon grundläggande mekanism som ger olika aspirineffekt vid olika födel-sedatum. Det förutsätts att ISIS-resultaten är slumpens spel i en statistiskt betänklig subgruppanalys av en randomiserad studie.

Svagheter hos den randomiserade studien

Den randomiserade studien har betytt oerhört mycket för en sund klinisk utveckling [3], men den har vissa inneboende

svagheter. De tyngst vägande av dessa förtecknas i det följande:

Hopklumpning av sjukdomar med olika patofysiologiska mekanismer. Om man under tidigt 1900-tal hade gjort en stor randomiserad studie på B₁₂ vid blodbrist hade resultatet troligen blivit positivt och B₁₂ följaktligen getts till alla patienter med blodbrist, vilket har påpekats av Maseri [4]. Den kunskap vi senare förvärvat om anemiens patogenes gör ju detta orimligt. På samma sätt kan man förmoda att många sjukdomsfall som studerats i moderna randomiserade studier har en mycket varierad patofysiologi, med konsekvensen att vissa patienter får behandling i onödan, eller till och med far illa av behandlingen.

En framtida detaljerad genetisk och patofysiologisk karakterisering av patienterna och deras sjukdomar kan förmodas ge en mer nyanserad syn på behandlingen än den som de randomiserade studierna gett. En sådan utveckling skulle minska läkemedelskostnaderna, som blir mycket stora när jättelika patientgrupper på basen av randomiserade studier skall genomgå kostsamma behandlingar.

FRISK II-undersökningen [5] har visat att subakut invasiv behandling med perkutan angioplastik/kranskärloperation är överlägsen non-invasiv behandling. Detta är en värdefull slutsats av stort praktiskt värde från en randomiserad studie. Den mycket snabba utvecklingen, med t ex förfinade sk stent vid perkutan angioplastik av kranskärlen, liksom raffinerade läkemedel av typ statiner, gör dock att resultaten kan vara föråldrade om några år. Balansen mellan effektivitet av invasiv respektive medikamentell behandling kan då mycket väl ha svängt. Det kommer att bli svårt att producera randomiserade kontrollerade studier som belyser aktuell teknisk standard och läkemedelsutveckling på ett så dynamiskt område som behandlingen av akut krans-kärlssjukdom.

Mängden av kliniska frågeställningar. Den kliniska vardagen innehåller mängder av beslut som aldrig kommer att kunna bli föremål för randomiserade studier. Det finns t ex många sådana studier om effekten av insättning av läkemedel, men mycket få studier som belyser effekten av utsättning av dessa läkemedel. Vår kunskap om hur länge medicinering bör pågå är därför ofta rudimentär. Läkemedelsbehandling av epilepsi är ett av de få undantagen från ovanstående.

Ingen kommer t ex att göra någon randomiserad kontrollerad studie av utsättning av antikoagulationsbehandling vid icke valvulärt förmaksflimmer hos multisyjuka människor i hög ålder. Läkemedelsstudier är nämligen vanligen kommersiellt finansierade. Frågeställning om medicinutsättning hör till den stora grupp problem i klinisk vardag, där klinikern utan hjälp av randomiserade studier måste söka sig fram till något slags rimlighetsbedömning.

Funderingar om Ansvarsnämndens reaktion om en patient får hjärnblödning under fortsatt antikoagulationsbehandling, respektive hjärninfarkt vid utsatt behandling, innebär säkert för många kliniker en stor osäkerhet i behandlingen av den beskrivna patientgruppen. Ofta får nog patienten fortsatt medicinering, på grund av att detta bedöms som medikojuridiskt säkrast.

Kombinerad behandling. Det är praktiskt omöjligt att testa många behandlingskombinationer i en randomiserad studie. Hos patienter med hjärninfarkt har det i sådana studier visat sig att acetylsalicylsyra, betablockerare, ACE-hämmare och statiner vardera minskar dödligheten med ca 20 procent, vilket rent matematiskt ger en osannolikt kraftig minskning av

morbidity och mortalitet hos dessa patienter. Det behövs egentligen en randomiserad studie som jämför alla tänkbara kombinationer av dessa fyra behandlingar, vilket är praktiskt omöjligt.

Optimal behandling inte detsamma som optimal vård

Optimal behandling är inte riktigt detsamma som optimal vård. I sistnämnda »vårdkunskap« måste vägas in individuella patientfaktorer, prioriteringsproblem och, inte minst, storleken på en behandlings fördel över en alternativ behandling. Denna fördel kan ju vara minimal. Någonstans går en rimlighetsgräns!

Läkarkonst förutsätter klinisk stolthet och ambition

Var det rätt av Bertil von Ahn och Eskil Kyhlin att utan stöd av randomiserade studier behandla på sätt som beskrivits? I backspegeln syns von Ahns handlande framsynt, men Kyhlins tveksamt. Episoderna illustrerar den praktiskt arbetande läkarens situation, som i grunden är ganska oförändrad. Denne måste göra en rimlig syntes av kliniska data, kunskap om patofysiologiska mekanismer, observationsstudier och eventuella randomiserade studier. Han måste också väga in personlig erfarenhet, som tyvärr ofta blir begränsad i en tid med mycket sammanträden, konferenser och medicinsk byråkrati.

Om detta praktiseras väl utövas läkarkonst. Denna konst får aldrig reduceras till förmågan att följa vissa instruktioner vilkas efterföljd granskas av kvalitetskontrollörer.

Läkarkonst förutsätter klinisk stolthet och ambition. Den kliniskt praktiserande läkaren står nära det okända gräns, och det är viktigt att göra bruk av nyfikenhet och iakttagelseförmåga. Dagens kliniska iakttagelser är grunden för framtidens evidensbaserade medicin. Goda beskrivningar av enstaka fall och begränsade patientserier är en mycket viktig del av medicinsk litteratur.

Stoikerna i det antika Grekland hävdade att den bästa bågskytten inte är den som träffar mitt i prick, utan den som utför bågskytten mest troget skyttekonstens regler [6]. Vindförhållanden och andra praktiska omständigheter kan nämligen göra att även den mest regeltrögne inte lyckas träffa målet, och detta bör inte ligga honom till last.

Det ligger en del i detta, men resonemanget blir på något sätt fel i sin ensidighet. Huvudsyftet är ju trots allt att träffa mitt i prick.

Översatt till medicinsk behandling leder ett renodlat »stoiskt« synsätt till en »guideline«-styrd evidensbaserad behandling. För ett verkligt gott behandlingsresultat fordras dock också att läkaren har sinne för det speciella i varje klinisk situation, god allmän människokänedom och intuition. Dessa ting utgör viktiga komplement till evidensbaserad medicin, komplement som är oundgängliga för god läkarkonst.

Referenser

1. FRISC Study Group. Low-molecular-weight heparin during instability in coronary artery disease. *Lancet* 1996; 347: 561-68.
2. ISIS-2 Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction. *Lancet* 1998; 2: 349-60.
3. Nilsson G, Sörensen S. Leve den randomiserade studien! *Läkartidningen* 2000; 97: 3445-9.
4. Maseri A. Evidence-based medicine: progress but not a final solution. *Clin Cardiol* 1998; 21: 163-4.
5. FRISC II investigators. Invasive compared with non-invasive treatment in unstable coronary-artery disease: FRISC II prospective randomised multicentre study. *Lancet* 1999; 354: 708-15.
6. Vandenbroucke JP. Evidence-based medicine and »Médecine d'Observation«. *J Clin Epidemiol* 1996; 49: 1335-8.

Jan-Otto Ottosson, professor emeritus, Göteborg (janotto.ottosson@swipnet.se)

Placeboeffekten behöver inte betvivlas!

II Per Daléns artikel i Läkartidningen 1–2/01 [1] ger två delvis motsägande budskap. För det första uttrycker Helsingforsdeklarationen entydigt att man inte bör jämföra en ny behandling med placebo utan med redan etablerade behandlingar. Kravet medför teoretiska och praktiska problem i behandlingsforskningen, som inte är föremålet i detta inlägg. För det andra ifrågasätts om placebo över huvud taget har någon effekt. I så fall skulle användningen av placebo inte motiveras av dess psykologiska effekter utan endast tjäna som kontroll av sjukdomens spontana förlopp och förväntningarna hos patienter och bedömare.

Beechers ytliga och ofta okritiskt citerade artikel [2] ger inte anledning att ifrågasätta existensen av placeboeffekter. Som Statens beredning för medicinsk utvärdering (SBU) har redovisat i rapporten om patient–läkarrelationen finns en övertygande dokumentation [3]. Placeboeffekter definieras här som effekter som uppkommer i interaktionen mellan patient och läkare genom att patienten tror på behandlingen och får sina förväntningar uppfyllda.

Drogen doktorn

Effekten av en god patient–läkarrelation är dokumenterad på allmänmedicinska och invärtesmedicinska mottagningar vid diabetes, hypertoni, magsårsliknande besvär, irritabel kolon samt diverse olika besvär.

Utrymmet medger endast referat av ett par av undersökningarna. Vid en allmänpraktik i Storbritannien ingick 200 patienter med hosta, halsont, förkylning eller värk i olika delar av kroppen utan objektiva fynd eller säker diagnos [4]. De fördelades slumpvis på fyra grupper: positivt bemötande med eller utan behandling och negativt bemötande med eller utan behandling. Det positiva bemötandet bestod i att patienten fick en diagnos och besked att det skulle vara bättre om några dagar. Om en medicin skrevs ut var det med försäkran om att den säkert skulle verka; om medicin inte skrevs ut motiverades det med att den inte behövdes. Det negativa bemötandet bestod i att läkaren förklarade att han inte var säker på vad som felades patienten. Om ett recept inte skrevs ut var det med motiveringen »och därför får du ingen behandling«; om recept skrevs ut var det med orden: »Jag är inte säker på att medicinen kommer att ha någon effekt.« Tabletter av B-vitamin (aneurin) användes som placebo.

Vid uppföljning efter två veckor hade 64 procent av de positivt bemötta patienterna blivit bättre, 39 procent av de nega-

SAMMANFATTAT

Placeboeffekten finns! Det är dokumenterat att ett positivt bemötande ger ett snabbare tillfrisknande än ett negativt. Ofta rör det sig om placebo förstärkning av specifikt verkande behandlingar.

Det är inte oetiskt att utnyttja placebo förstärkning av behandlingar. Att däremot ge placebo utan informerat samtycke, eller att ge specifika behandlingar enbart för att dra nytta av deras placeboeffekter, är olika former av bedrägeri som riskerar att äventyra patienters förtroende för vården.

Det vore mer konstruktivt att bedöma patient–läkarrelationens inverkan på vårdens resultat än att enbart uppfatta placebo som en störande faktor vid behandlingsvärdering.

Evidensbaserad medicin

tivt bemötta. Det var ingen skillnad om behandling hade givits eller inte (53 procent mot 50 procent förbättrade). Försöket visar att sättet att bemöta patienterna var avgörande för förloppet. Genom sitt bemötande skapar läkaren antingen positiva förväntningar om ett gott resultat eller negativa förväntningar om ett dåligt resultat av behandlingen.

Placebo förstärkning

En annan belysande studie utfördes vid en svensk vårdcentral där effekten av att läkaren avsiktligt maximerade förväntans effekten av penicillinbehandling vid halsfluss undersöktes [5]. Undersökningen utgick från 100 patienter som av mottagningsköterska bedömdes ha akut tonsillit. Patienterna fördelades slumpvis på två grupper och fick möta samma läkare. Sedan läkaren presenterat sig undersöktes experimentgruppens patienter (n=50) först i mörkrum med specialbelysning och togs sedan till läkarens rum, där utförlig information gavs om diagnos, behandling och prognos. Läkaren betonade att

patienten troligen skulle må bra efter 24 timmar. Patienten erbjöds att ställa ytterligare frågor. Konsultationen tog omkring tio minuter. För kontrollgruppens patienter (n=50) presenterade sig inte läkaren, patienterna genomgick en mer summarisk undersökning på läkarens rum och fick ingen information om prognosen. Konsultationen tog omkring sex minuter. Från samtliga patienter togs prov för svalgodling. Alla fick också recept på penicillintabletter och sjukskrevs om de begärde det. De meddelades också att de skulle kontaktas per telefon efter två dagar.

Åldersfördelningen var likartad i grupperna. Som väntat, med tanke på exponering från barn i daghem och skolor, var det en övertikt av kvinnor (65 procent). Flertalet hade sökt vård inom en vecka efter symtomdebuten. Betahemolyserande streptokocker förekom hos 58 patienter, av vilka 32 slumpvis hade hamnat i experimentgruppen och 26 i kontrollgruppen. Flertalet av de övriga hade sannolikt virusinfektioner.

Vid efterundersökningen ställdes i förväg formulerade frågor. Intervjuaren var blind för patienternas grupptillhörighet. Det visade sig att fler i experimentgruppen än i kontrollgruppen hade känt att symtomen från halsen hade förbättrats och att behandlingen hade hjälpt dem. Fler ansåg också att de hade fått tillräcklig information om sin sjukdom och behandling. Skillnaderna fanns vid båda infektionstyperna men var störst i streptokockgruppen

Sättet att bemöta patienterna hade i denna undersökning inneburit en tilläggs effekt till penicillinbehandlingen hos patienterna med streptokockinfektion. Patienterna med virusinfektioner hade ingen nytta av penicillinbehandlingen, men placebobehandlingen hade givit en positiv effekt. Det är fullt möjligt att det man tror är en specifik behandling i själva verket är en placeboeffekt. Man kan med andra ord få en placeboeffekt utan att ge placebo.

Sanna behandlings- och placeboeffekter

Det som i många sammanhang kallas placeboeffekter är en kombination av en sann placeboeffekt, naturalförlopp, regression mot medeltalet, andra tidsrelaterade effekter, oidentifierade parallella interventioner och systematiska fel, »bias«. Eftersom allt detta kontrolleras vid randomiserade kontrollerade prövningar utgör en sådan vid definition inget hinder mot uppskattning av sanna behandlingseffekter. För att få en uppfattning om en sann placeboeffekt fordras däremot en obehandlad kontrollgrupp [6].

Placeboeffekter på smärta

En sökning i Medline av jämförelser mellan placebobehandling och ingen behandling alls avkastade tolv studier, varav sju gällde smärta [6]. Sammantaget visade sex av smärtestudierna att en placeboeffekt förelåg. I den enda studien utan skillnad från kontrollgruppen hade placebo givits postoperativt till fortfarande sovande patienter. Mer handfasta placebo-behandlingar, som simulerat ultraljud eller låtsad akupunktur, gav kraftigare effekter än orala placebo. Objektiva fynd såsom minskad svullnad och ändrad koncentration av C-reaktivt protein ingick bland placeboeffekterna.

Dubbelblinda jämförelser mellan randomiserade placebo-behandlade och obehandlade patientgrupper kontrollerar för alla skenbara placeboeffekter utom patienternas egna förväntningar, vilka dock utgör en del av placeboeffekten.

Placebobehandlingars etik

Eftersom alla åtgärder som bidrar till patienters snabba förbättring är önskvärda är det inte oetiskt att utnyttja en placeboeffekt, så mycket mer som den endast innebär ett tillgodo-seende av patienternas lagstadgade rätt att få information och bemötas med omtanke och respekt. En placeboförstärkning

-IDAG MÅR JAG
FINT MEN DET
ÄR NOG BARA
PLACEBOEFFEKT!



ILLUSTRATION: LASSE PERSSON

av specifikt verkande behandlingar uppstår genom patientens och läkarens gemensamma förväntningar på att en behandling skall visa sig framgångsrik. Dessa förväntningar får i sin tur sin näring av ritualer i samband med behandlingen, en god patient-läkarrelation och allt annat som ingår i den totala vårdmiljön. Detta är okontroversiellt från etisk synpunkt.

I övrigt bör grunden i all sjukvård och forskning vara att patient och läkare kan lita på varandra. Placebo i form av en överksam tablett eller injektion kan därför inte användas för att fastställa att en patients besvär är »verkliga« om placebo-effekten är negativ, eller avfärda dem som »inbillade« om placeboeffekten är positiv. Att i diagnostiskt eller terapeutiskt syfte ge placebo utan informerat samtycke är en form av bedrägeri som, när det avslöjas, riskerar att äventyra patient-läkarrelationen och patientens förtroende för sjukvården överhuvudtaget.

Etiskt betänkligt är det också att ge behandlingar som inte har några specifika nyttoeffekter, utan enbart för att dra nytta av deras placeboeffekter. När läkaren är medveten om frånvaron av specifika effekter av en behandling, men låter patienten förbli okunnig härom, kan det återigen ses som en form av bedrägeri. Medicinens historia ger exempel på blindtarmar och tonsiller som utan klara indikationer opererats bort hos barn endast för att ge intryck av att kraftfulla åtgärder har vidtagits eller för att föräldrarna trycker på. Borttagande av livmodern och andra operationer har gjorts på vuxna av likartade skäl. Onödiga diagnostiska undersökningar har utförts för att patienterna skall få intrycket att man verkligen anstränger sig för deras skull. Det har beräknats att 35–45 procent av alla recept gäller läkemedel som inte har någon verkan på det tillstånd för vilket de förskrivs [3]. Att de ändå skrivs ut beror på en olycklig kombination av otåliga patienter, jäktade läkare och en aktiv läkemedelsreklam.

I alla sådana fall får placeboeffekten köpas till priset av obehag och risker. Patienter, försäkringsbolag och det allmänna belastas av utgifter, allergi kan utvecklas mot mediciner och resistens kan uppstå mot antibiotika när dessa ges vid virusinfektioner. När de rätta förhållandena avslöjas riskerar

sjukvården att drabbas av en förtroendekris som går ut över framtida patienter.

Standardisering av försöksbetingelser

Rimligen blir placeboeffekten optimal i en god patient-läkarrelation. I sitt depressionsforskningsprogram har National Institute of Mental Health framhållit att en god relation inrymmer förtroende, empati, stöd, optimistisk attityd, information om sjukdomen och behandlingen, råd rörande sociala förhållningssätt och avreagering av känslor [7, 8]. Som renodlad psykoterapi räknas däremot fokusering på speciella interpersonella förhållanden, tolkningar, klarifikationer av känslor mot terapeuten eller andra, beteendestruktioner som går utöver enkla råd, psykodynamiska förklaringar och interpersonell psykoterapi. Vid psykoterapi består placeboeffekten av den sk terapeutiska alliansen som, oberoende av terapiform, är den faktor som starkast bidrar till terapins resultat [9].

Placebotermen ifrågasatt

Termen placebo ger negativa associationer till en störande faktor vid behandlingsvärdering. Med hänsyn till de många positiva effekterna har föreslagits att termen helt bör utgå [10, 11]. I stället bör man sträva efter att bedöma inflytandet av alla de faktorer som utöver specifika behandlingseffekter styr vårdresultatet: naturalförlopp, regression mot medeltalet, patientens och läkarens förväntningar, patient-läkarrelationens kvalitet, behandlingssituation, betingning och grad av ångest.

Referenser

1. Dalén P. Rutinmässig användning av placebo ifrågasätts. Tydligt ställningstagande i Helsingforsdeklarationen. Läkartidningen 2001; 98:16-8.
2. Beecher HK. The powerful placebo. JAMA 1955;159:1602-6.
3. Ottosson JO, red. Patient-läkarrelationen. Läkekonst på vetenskaplig grund. Stockholm: SBU och Natur och Kultur, 1999.
4. Thomas KB. General practice consultations: is there any point in being positive? BMJ 1987;294:1200-2.
5. Olsson B, Olsson B, Tibblin G. Effects of patients' expectations on recovery from acute tonsillitis. Fam Pract 1989;6:188-92.
6. Ernst E, Resch KL. Concept of true and perceived placebo effects. BMJ 1995;311:551-3.
7. Fawcett J, Epstein P, Fiester SJ, Elkin I, Autry JH. Clinical management - imipramine/placebo administration manual. NIMH treatment of depression collaborative research program. Psychopharmacol Bull 1987;23: 309-21.
8. Malt UF, Robak OH, Madsbu HP, Bakke O, Loeb M. The Norwegian naturalistic treatment study of depression in general practice (NORDEP). I: randomized double blind study. BMJ 1999; 318: 1180-4.
9. Horvath AO, Symonds RD. Relation between working alliance and outcome in psychotherapy: A meta-analysis. Journal of Counselling Psychology 1991; 38: 139-49.
10. Roberts AH. »The powerful placebo« revisited. What power? What placebo? Presented at the seventh world congress of the international association for the study of pain; August 23, 1993. Paris, France.
11. Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. JAMA 1994;271:1609-14.

Särtryck

Läkartidningen

Alla kroppens celler reagerar på olika signalämnen i omgivningen, ämnen som styr deras fundamentala livsprocesser.

Dessa ämnen kallas kollektivt tillväxtfaktorer. En serie i Läkartidningen 1995 om dem speglar tendenser i dagens medicinska forskning och pekar på några tillämpningsområden.

Området är i början av en snabb utveckling och många produkter är under utprovning för klinisk användning.

Häftet omfattar 12 artiklar på sammanlagt 56 sidor + färgomslag.

Priset är 90 kronor.



Tillväxtfaktorer

Beställer härmed.....ex
av "Tillväxtfaktorer"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till Läkartidningen
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Evidensbaserad medicin:

Tio kritiska invändningar mot en värdefull metod

Evidensbaserad medicin (EBM) innebär en kunskapsbaserad, systematisk granskning av medicinsk vetenskap med relevans för klinisk praxis. Begreppet har det senaste årtiondet kommit att bli centralt inom praktisk medicin, och även vidgats till evidensbaserad omvårdnad [1].

CARL-MAGNUS STOLT
chefläkare Södra Älvsborgs sjukhus, Borås; professor i humanistisk medicin, Karolinska institutet, Stockholm
Carl-Magnus.Stolt@his.ki.se

II Det rör sig om en framväxande internationell rörelse inom medicinen med syfte att ge bästa tänkbara beslutsunderlag. Detta är nödvändigt i det alltmer massiva informationsflödet. Ingen enskild läkare kan längre överblicka och värdera kunskapsstoffet ens inom ett tämligen begränsat terapeutiskt område.

Risken finns då att en patient utsätts för en mindre bra behandling. Detta kan vara ett hot mot patientsäkerheten. Istället för att faktabasera den vård man erbjuder, kanske läkaren ägnar sig åt en eller flera av följande alternativa baser för sin praktik: en auktoritetsbaserad, en emotionellt baserad, en provinsialt baserad, en asteniskt baserad, eller kanske i vissa akuta situationer en nervöst baserad medicin!

EBMs styrka ligger således i att den enskilda läkaren får en färsk, uppdaterad och oberoende kunskap, förhoppningsvis fri från exempelvis läkemedelsindustrins inflytande. Kunskapen blir därmed mer oberoende av tyckande; dess subjektivitet begränsas.

Störst tyngd i utvärderingen av vetenskapliga uppsatser ges prospektiva randomiserade studier. Det evidensbaserade kunskapsunderlaget är också tillgängligt för alla, såväl läkare som allmänhet. Att det finns en relativ konsensus om vad som är en bra behandling betyder också att olika läkare lämnar ungefär samma budskap till en patient i en given behandlingssituation, exempelvis vad gäller effektiviteten av olika behandlingar vid ryggvärk.

Granska granskaren!

EBM brukar betraktas som en kritisk granskning av medicinen, och inte minst därför finns det anledning att kritiskt

granska även denna! Ett förutsättningslöst inommedicinskt, vetenskapsteoretiskt kritiskt förhållningssätt säger att allt måste kunna granskas!

Finns det svagheter i EBM-begreppet? Jag presenterar här tio kritiska invändningar av olika bärighet.

Det nya hinner inte bli evidensbaserat

1. Den första invändningen, eller begränsningen, är att det riktigt nya inte kan hinna bli evidensbaserat. Det finns alltså en relativ tröghet i EBM-metoden, men denna tröghet kan givetvis även vara kvalitetsbevarande!

2. Det finns vidare en begränsning i vad som är undersökt, och denna begränsning vidlåder naturligt nog även EBM.

3. En tredje begränsning är att en stor del av den forskning som granskas är kommersiellt styrd.

Risk för standardiserad kokboksmedicin

4. Den fjärde invändningen gäller risken för överdriven standardisering, vilket medför en konformism som likriktar och riskerar att hämma den medicinske praktikern. Med denna förnämning av den medicinska praktiken riskerar vi att få en »kokboksmedicin«, som åsidosätter verklighetens mångfald och mångtydighet [2].

Man ser då inte patienten i sitt sammanhang. Man ägnar sig mer åt en »medianbaserad« medicin, när man kanske istället borde ha ett inslag av personorienterad dialogisk »narrative-based« medicin [3]. Man kan ställa sig frågan: Studiepatienter, finns de i verkligheten? En populationsbaserad kunskap är uppenbart inte alltid möjlig att överföra till den enskilda individen.

Subjektivt urval av studier

5. En femte kritisk punkt vid granskningen av EBM är att urvalet av studier man väljer att EBM-granska i slutändan alltid är subjektivt. Man väljer ur ett be-

tydligt större underlag av olika anledningar ut ett antal uppsatser till sin evidensbaserad. Vad styr urvalet? Det finns i det sammanhanget också en risk att man sätter just randomiserade kontrollerade prospektiva studier på en alltför hög piedestal. Medicinens historia lär oss att många betydande bidrag till medicinens möjligheter tillkommit på annat vis!

Hälsoekonomins inflytande

6. En sjätte invändning mot EBM-begreppet är den fara som finns om man alltför hårt kopplar begreppet till hälsoekonomiska resonemang. Då hotar ekonomin att alltför mycket styra synen på vad som är den bästa behandlingen – eller det bästa behandlingsunderlaget.

Jag påstår ingalunda att sjukvårdens ekonomiska dimension är oviktig, men den kommer in i ett senare skede, inte i det strikt medicinska kunskapsunderlaget.

Detta är måhända en kontroversiell invändning! Kopplingen mellan EBM och hälsoekonomi har tonats ned avsevärt de senaste åren, inte minst i Sverige. Men ännu 1995 kunde man i tongivande internationella EBM-sammanhang läsa att målsättningen med EBM var att eliminera dyra, ineffektiva och farliga be-

EBM är en värdefull metod för att skaffa så bra oberoende beslutsunderlag som möjligt. Men liksom allt annat i medicinen måste det också tåla kritisk granskning, och vi måste resonera om och inse dess begränsningar!

handlingar [4]. Ineffektiva och farliga är lätt att instämna i, men dyra?

Relationen patient–läkare i skymundan

7. Den sjunde invändningen är den jag anser vara den tyngst vägande. I en strikt EBM-baserad medicinsk praktik riskerar viktiga aspekter av patient–läkarrelationen att hamna i skymundan. Orsaken till detta måste förstås från insikten i att medicinsk praktik består av en kombination av medicinsk vetenskap och läkekonst.

Det som ryms i begreppet läkekonst låter sig inte så lätt evidensbaseras. I en förnämlig SBU-bok [5] granskade man läkekonst på vetenskaplig grund. För att förstå problemet måste man först vidga den traditionella medicinens metodologi till att, förutom kvantitativ metod, även inbegripa kvalitativ sådan.

Kan tröst vetenskapliggöras?

Ytterst är detta en fråga om vad som är vetenskap. Vi har som medicinare ofta en inställning att endast naturvetenskap är vetenskap. Det är naturligtvis en försnävad uppfattning. Skall man studera och vetenskapliggöra det som ryms i läkekonsten måste man använda såväl kvantitativa som kvalitativa metoder.

OVANA VID KVALITATIVA STUDIER?

I EBM-sammanhang har främst kvantitativa, naturvetenskapliga studier varit i blickfånget, och det är de som ges störst tyngd. På senare år har detta förändrats och det är positivt. SBU-boken om läkekonst tog även upp kvalitativa studier till granskning, men de kvantitativa vägde tveklöst tyngst i slutsatserna. Detta kan bero på att de kvalitativa studierna helt enkelt var sämre, men det kan också bero på en ovana vid att bedöma kvalitativ forskning.

I denna min sjunde invändning mot EBM-begreppet – att läkekonsten kommer i skymundan – finns ännu en dimension. Man kan nämligen fråga sig om det till och med finns företeelser i

läkekonsten som inte ens kan vetenskapliggöras, alldeles oavsett val av metod. Centralt i läkekonsten finns trösten. Att vetenskapliggöra trösten går nog, men det är inte många som försökt!

Kanske är det symtomatiskt att ordet tröst inte nämns en enda gång i den ovan nämnda SBU-boken. Låt mig i sammanhanget ändå säga att jag ändå anser den aktuella boken oerhört värdefull, trots denna begränsning [6].

Det som kanske inte kan utforskas – är det inget vi skall bry oss om? Det som undflyr och gäckar vetenskapliggörande – är det oväsentligt? Låt mig bara konstatera att på detta område kan konsten – t ex skönlitteraturen – komplettera den medicinska vetenskapen.

Kvantitativ förtjuning i oväsentligheter?

8. En åttonde invändning har med vetenskaplig metod att göra, och har beröringspunkter med föregående invändning. Det handlar om risken att övervärdera kvantitativ metod, och detta problem har också förtjänstfullt uppmärksamats av företrädare för Statens be-

redning för medicinsk utvärdering (SBU). Man har självkritiskt resonerat om »surrogat-effektmått«, när man mäter något som inte betyder något för patienten [7].

Bortser vi från hälsoperspektivet?

9. Den nionde invändningen rör en kritik som ofta riktas mot medicinen rent allmänt, inte minst från exempelvis vårdvetare och kulturvetare med intresse för medicin och vård. Det gäller risken att vi inom medicinen fixerar sjukdomen och bortser från hälsoperspektivet. Med EBM förstärks kanske denna risk. Vi måste minnas Hippokrates ord om att det egentligen inte finns några sjukdomar, utan bara människor som kan bli sjuka. Den levande människan har i sig alltid ett visst mått av hälsa, som inte sällan kan förstärkas. Ett sätt att förstärka hälsa är förstås att bota och lindra sjukdom.

Tål EBM en EBM-granskning?

10. Den tionde och sista kritiska tanken inför EBM-begreppet är mycket övergripande. Är EBM i sig evidensbaserad? Tål EBM att granskas av sig själv? Blir folk friskare om vi tillämpar EBM?

Sammanfattning

Avslutningsvis kan jag konstatera att EBM är en värdefull metod för att skaffa så bra oberoende beslutsunderlag som möjligt. Men liksom allt annat i medicinen måste det också tåla kritisk granskning, och vi måste resonera om och inse dess begränsningar!

Referenser

1. Eriksson K, Nordman T, Myllymäki I. Den trojanska hästen. Evidensbaserat vårdande och vårdarbete ur ett vårdvetenskapligt perspektiv. Åbo: Institutionen för vårdvetenskap, Åbo akademi, Vasa, 1999.
2. Tonelli MR. The philosophical limits of evidence-based medicine. *Academic Medicine* 1998;73:1234-40.
3. Greenhalgh T, Hurwitz B, eds. Narrative based medicine. Dialogue and discourse in clinical practice. London: BMJ Books, 1998.
4. Rosenberg W, Donald A. Evidence-based medicine: an approach to clinical problem-solving. *BMJ* 1995;310:1122-6.
5. Ottoson JO, red. Patient–läkarrelationen. Läkekonst på vetenskaplig grund. Stockholm: Natur och Kultur i samarbete med SBU, 1999.
6. Stolt CM. Viktig bok om medicinen som konst och vetenskap. *Läkartidningen* 1999; 97:1312-3.
7. Eliasson M. Evidensbaserad medicin »begins at home«. *Läkartidningen* 2000;98: 3860-5.

Med facit i hand kan man självklart alltid hävda att beslut att avstå från hjärt-lungräddning var felaktigt i de fall patienten tillfrisknar. Frågan är emellertid om det i dessa fall är rimligt att utdela disciplinpåföljd till de läkare som deltagit i sådana beslut? Effekten av detta kan för framtiden bli att dessa mycket viktiga beslut inte dokumenteras eller att man helt enkelt avstår från att fatta beslut om begränsning i behandlingen.

Hemlighetsfulla koder

Förhoppningsvis är den tid förbi då beslut om att avstå från hjärt-lungräddning noterades med lokala koder någonstans i journalen i form av symboler eller bokstavskombinationer vars innebörd bara var känd för berörd sjukvårdspersonal.

Det är snarast så att dessa beslut måste diskuteras i en öppen atmosfär, där besluten är tydliga för alla inblandade – patienter, anhöriga, läkare och övrig sjukvårdspersonal. Det får inte bli så att enskilda läkare kommer att undvika att ta ställning i svåra etiska frågor som dessa av rädsla för att bli anmälda till HSAN.

I den aktuella situationen hade patienten tre veckor efter intoxicationen kända strukturella hjärnskador bilateralt i talamus, mediala temporalloberna och hippocampi. Beslutet att ej återuppliva patienten vid ett eventuellt cirkulations- eller andningsstillestånd förefaller inte orimligt. Någon annan förändring av patientens övriga vård skedde så vitt känt inte.

Det förefaller ogenomtänkt från HSAN att ett väl övervägt medicinskt beslut, som dessutom inte haft någon verklig påverkan på vårdförloppet, ska leda till disciplinpåföljd. Däremot är det mycket väsentligt att dessa beslut noggrant dokumenteras.

Referenser

1. Sjökvist P, Sundin PO, Berggren L. Limiting life support. Experiences with a special protocol. *Acta Anesthesiol Scand* 1998; 42:232-7.
2. Melltorp G, Nilstun T. Decisions to forego life-sustaining treatment and the duty of documentation. *Intensive Care Med* 1996; 22:1015-9.

Evidensbaserad medicin:

Många kritiker missförstår EBM

När man diskuterar ett sådant fenomen som Evidence Based Medicine (EBM) är det viktigt att man är överens om vad man talar om. Många som är kritiska mot företeelsen EBM missförstår vad det är fråga om. Detta framkommer också i det inlägg som Carl-Magnus Stolt gjorde i LT 28-29 (sidan 3255).

LARS WERKÖ
professor, Stockholm

■ Jag kommenterar nedan några av Stolts punkter.

Men först:

Den väsentliga utgångspunkten för hela begreppet är att man skall veta exakt *på vilken evidens* man fattar beslut i sjukvården, inte att man endast kan använda strikt vetenskapligt framtagna data som underlag för beslut.

Alla inser att det finns olika grader av evidens för olika kunskaper, vilket också framgår av Stolts inlägg. I den evidensbaserade vården gäller det att veta på vilken evidens de metoder vilar som jag vill utnyttja. Jag skall använda dem som vilar på bästa möjliga evidens, och för att veta detta måste jag följa den aktuella litteraturen.

Den kliniska verkligheten

1. Invändningen att det riktigt nya inte har hunnit bli evidensbaserat baseras på att Stolt vet att detta nya är effektivt, inte bara nytt. För att veta att det är effektivt måste det studeras, och den kliniske forskaren som är angelägen om att införa nya metoder har skyldighet att själv delta i vetenskapliga studier för att försäkra sig om metodens effektivitet om ingen annan har gjort det.

Införs nya metoder innan de är kliniskt utprovade förlänger man tiden innan man vet deras värde. Eventuellt minskas möjligheten att göra kontrollerade studier – se laparoskopisk kirurgi. Egentligen innebär EBM ett åtagande att stödja klinisk, patientorienterad forskning. Det är ju endast i den kliniska verkligheten, inte i det teoretiska laboratoriet, som effektivitet kan bevisas.

2. Begränsningen av vad som är undersökt är inte något argument mot EBM. Man skall bara veta vilka begränsningar metoden man vill använda har i fråga om hur väl studerad den är.

3. Så mycket större anledning är det att granska underlaget om det endast är kommersiellt styrt.

Ej »kokboksmedicin«

4. EBM innebär inte att man måste följa sådana riktlinjer som leder till »kokboksmedicin«. Tvärtom – man är skyldig att själv granska det vetenskapliga underlaget, inklusive det patientmaterial det är baserat på i förhållande till det aktuella problemet med den patient man har framför sig. Riktlinjer framtagna av så kallade experter är värdelösa om de inte är baserade på kritisk granskning av aktuell litteratur. Det finns alltför många sådana.

5 och 6. Är teknikaliteter som vidlå-

der alla undersökningar i detta sammanhang

6. Detta är den viktigaste punkten i kritiken, men även den bortser från att det alltid, även när det gäller att arbeta enligt EBM, är den enskilde läkaren som tillsammans med patienten bestämmer. Den enskilde patienten är unik och det kan alltid finnas omständigheter som gör att vad som kan tyckas vara klart bevisat i vetenskapliga studier ändå inte kan tillämpas just nu.

I ett långsiktigt arbete i ett mindre sjukhus i norra Italien, där läkarna varit angelägna att lära sig att använda EBM, har man gjort så att i varje fall, läkaren först, grundat på sin kunskap om patienten, den erfarenhet han själv haft i liknande fall och tidigare kunskap, lägger upp en plan för diagnos och behandling. Han går sedan till litteraturen och tar fram den evidens för hur liknande fall bör tas om hand, och gör tillsammans med patienten en syntes av sitt eget kunnande, vad litteraturen visar och vad patienten vill efter att ha blivit presenterad alla fakta.

7. Det största misstaget i fråga om tillämpning av EBM i praktisk sjukvård är att tro att den enskilde läkaren och den individuella patienten skulle vara bortkopplade från beslutet om utredning och behandling.

8. Självklart skall EBM granskas och diskuteras. I sina rapporter granskar SBU systematiskt det vetenskapliga underlaget. Läkartidningen har också på ett föredömligt sätt uppmuntrat till det.

Komplicerad verklighet

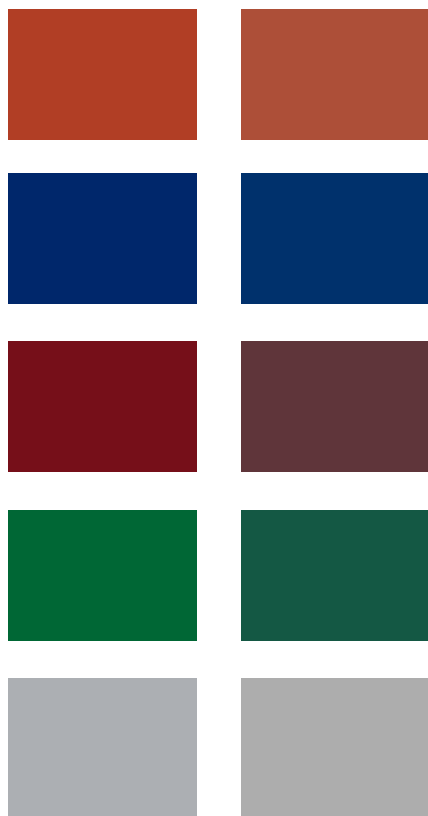
Ytterligare en kommentar. Det har upp-

stått ett förklarligt missförstånd genom att EBM har kommit att diskuteras samtidigt som Cochrane Collaboration har etablerats och med stor framgång granskar den litteratur på vilken vi baserar våra åtgärder och beslut. Man skall då ha klart för sig att Cochrane Collaboration i huvudsak begränsar sin verksamhet till resultat av kontrollerade randomiserade studier, medan de som försvar sig till EBM måste göra det bästa möjliga i den komplicerade verkligheten som klinisk medicin utgör.

Att arbeta enligt EBM betyder således inte att man endast godkänner resultat av kontrollerade studier, utan att man vet när det finns sådana – och, när det inte finns sådana, vet hur väl en viss metod är undersökt. För att praktisera enligt EBM skall man alltså vara bekant med var gränsen går i aktuell vetenskaplig kunskap och använda bästa möjliga kunskap som underlag för sitt handlande. •



Test av färger



Välkommen debatt om evidensbaserad medicin

Det är glädjande att flera skribenter antagit min utmaning att formulera kritiska invändningar mot EBM-begrepp, först Göran Nilsson i nr 20 och nu Carl-Magnus Stolt i nr 28–29. Endast genom en öppen och ifrågasättande debatt kan vi undvika att EBM »ikoniseras« och blir till ett lika auktoritärt begrepp som den »opion-based medicine« som vi vill lägga bakom oss.

MATS ELIASSON
docent, sakkunnig i invärtes medicin,
SBU; medicinsk redaktör,
Läkartidningen
Mats.Eliasson@nll.se

II Orealistiska förväntningar om EBM:s möjlighet att lösa mångfalden av problem som vården står inför riskerar att leda till en »back-lash« med uppgivenhet kring möjligheten att använda bästa vetenskap som grund för klinisk verksamhet.

Lars Werkö har i ovanstående inlägg bemött de flesta av Stolts synpunkter men vissa bör kommenteras ytterligare. Flera av invändningarna kommer att beröras i de artiklar som under hösten avslutar Läkartidningens EBM-satsning.

Transparent process

1. Carl-Magnus Stolt är bekymrad över att urvalet av studier för »EBM-granskning« är subjektivt. Kommentaren är förvånande då subjektiviteten vanligen får ett betydligt större utrymme i andra sammanhang än de systematiska översikter som utgör en av hörnstenarna i EBM.

Urvalet baseras på explicita och a priori uppsatta kriterier, och hela processen är transparent för läsaren till skillnad från traditionella översiktsartiklar där läsaren helt är utelämnad avseende urval och tolkning. Användandet av två oberoende granskare reducerar också subjektiviteten avsevärt.

Kvalitativ forskning

2. Brist på användande av kvalitativa studier inom EBM bekymrar Stolt. Kvalitativa studier besvarar oftast frågor om hur olika aspekter på sjukdom och hälsa upplevs, medan EBM-rörelsen har tagit sin utgångspunkt i frågan om vilka medicinska metoder som har en gynnsam effekt för patienter på överlevnad och liknande.

Den kvalitativa forskningen kan dock bidra med viktiga pusselbitar. (Varför tar patienten inte medicinen? Hur resonerar patienten som ställs inför valet av operation av sin prostatacancer eller observation?) Läkartidningens EBM-serie

har därför i artiklar av Povl Riis och Thörn, Bunne och Hallberg i nr 12/2001 utvecklat rollen av sådan forskning inom en evidensbaserad sjukvård.

SBU och deras internationella systerorganisationer följer denna forskning med intresse, men det är uppenbart att möjligheten att kritiskt granska och validera resultaten för närvarande är begränsad och att dessa metoder behöver utvecklas ytterligare.

Surrogatmått ifrågasätts

3. Den mest svårbegripliga invändningen är risken med att övervärdera kvantitativ metod, och Stolt formulerar att SBU har »självkritiskt resonerat om surrogatmått«. Tvärtom så har ju just EBM-rörelsen gått i bräschen för att ifrågasätta surrogatmått såsom PEF-värden, CD-4-nivåer, radiologisk regress av tumörer osv.

De senaste årens SBU-rapporter har varit pionjärarbeten i att välja ut effektmått med direkt betydelse för patienten som grund för sina slutsatser. T ex kan rapporten om behandling av astma och KOL från år 2000 ses som ett genombrott för att använda livskvalitet som mått på en behandlings effekt och nytta för patienten.

Utan EBM-rörelsens ifrågasättande hade vi fortsatt i ett trask av tveksamma intermediärvariabler som läkemedelsrepresentanterna visar upp på färgglada OH-bilder. Frågan om surrogatmått belyses ytterligare i en kommande artikel.

4. Att EBM skulle fixera ett sjukdomsperspektiv är väl att skjuta på budbäraren? Ett evidensbaserat synsätt styr knappast frågan om synen på friskt eller sjukt men kräver den vetenskapliga basen synliggjord, vilket man dessutom gärna skulle vilja se mer av vad gäller den s k salutogenesforskningen.

Hjälpmedel – ej ett mantra

EBM är inget mantra utan ett hjälpmedel i klinisk vardag men för god läkekonst krävs ytterligare färdigheter och kunskaper.

Vi välkomnar fler inlägg från läsarna hur den goda vetenskapen leder till den goda vården! •

Kjell Asplund, professor, överläkare, medicinkliniken, Norrlands universitetssjukhus, Umeå, styrelseordförande, SBU

Den evidensbaserade medicinen är nödvändig men inte tillräcklig

Bör kompletteras inom områden där det vetenskapliga underlaget är svagt

II Läkartidningens artikelserie om evidensbaserad medicin (EBM) har ingående diskuterat fördelarna med ett evidensbaserat kliniskt handlande. En nyckelkomponent i EBM är den kritiska ansatsen. Det måste gälla också EBM i sig – detta paradig måste tåla att granskas, vägas, diskuteras. Denna artikel om samspelet mellan EBM och klinisk praxis är skriven i den andan och kompletterar den som Carl-Magnus Stolt, professor i humanistisk medicin, nyligen publicerade i Läkartidningen med tio invändningar mot EBM [1]. Den som söker en mer engagerad redogörelse för EBMs alla förtjänster uppmanas läsa de tidigare artiklarna i denna serie.

Det kliniska tänkandet

EBM framhåller det deduktiva tänkandet: ett kliniskt beslut nås genom att dra slutsatser utifrån bästa tillgängliga vetenskapliga underlag [2]. Studier av hur erfarna kliniker når fram till beslut har i stället lyft fram deras förmåga till analogt tänkande. Detta analoga tänkande bygger i hög grad på intuitiva referenser till andra patienter läkaren mött, på likheter och skillnader i fråga om symtompresentation, emotionell framtoning, fysikaliska fynd, laboratorieresultat etc [3].

Onekligan ligger EBMs deduktiva process närmare den klassiska vetenskapliga processen än vad klinikers ofta analoga tänkande gör (även om riktigt nyskapande forskning ofta bygger på intuitiva, analoga tankar). God läkekonst kräver plats för både det deduktiva och det analoga tänkandet.

Vetenskapligt underlag och klinisk praxis

Det finns ett intrikat samspel mellan den kliniska medicinen, individinriktad, och den kliniska forskningen, oftast gruppriktad. En av forskningens centrala uppgifter är att söka det allmängiltiga. Den måste därför ofta eliminera, eller åtminstone minimera, de individuella variationerna. Men i mötet med den enskilda patienten är det personliga, individuella perspektivet en central komponent.

När man håller sig strikt till det vetenskapliga underlaget, kan man lova en optimal chans till ett gynnsamt utfall, det kan gälla risken att insjukna eller dö, bästa chans till god funktion eller god livskvalitet. Men denna statistiskt baserade chans är inte detsamma som ett löfte om gott utfall i en enskild klinisk situation [4]. Den goda klinikern förmår att integrera det individuella perspektivet med den vetenskapliga basen. De flesta erfarna läkare kan nog erinra sig situationer där de frångått

SAMMANFATTAT

Evidensbaserad medicin är en nödvändig men inte tillräcklig komponent i god läkekonst.

EBM stärker en av de viktiga komponenterna i läkekonsten, den vetenskapliga grunden för vårt kliniska handlande. Men det finns en risk att andra värden skymms.

EBM grundar sig på studier av större patientgrupper. I mötet mellan patient och läkare är det i stället ofta de individuella variationerna som är avgörande.

EBM kan, på gott och ont, innebära att fokus förflyttas från det som är till gagn för den individuella patienten till samhällsnyttan.

Inom områden där det vetenskapliga underlaget är svagt, t ex omvårdnad, rehabilitering och sjukvårdsorganisation, är EBM till föga hjälp, särskilt inte som underlag för prioriteringar.

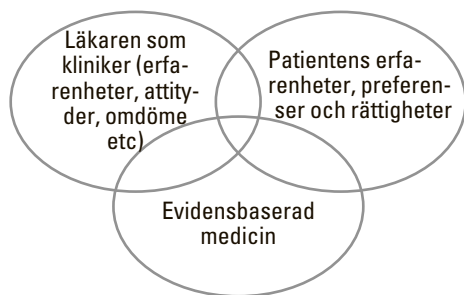
EBM måste kompletteras med strategier för att hantera kliniska situationer där ett vetenskapligt underlag saknas.

Evidensbaserad medicin

vad kliniska riktlinjer och vetenskapligt underlag säger och att detta visat sig vara klokt.

Hur hantera osäkerheten?

Problemet kan formuleras på följande sätt: Vem vill avvika med risk att betraktas som ovetenskaplig? Att försvara ett kliniskt beslut med utgångspunkt i den egna kliniska erfarenheten har blivit till något av ett kastmärke för den gammalmodiga, auktoritära läkaren [5]. Att strikt följa alla riktlinjer är lätt att försvara, att göra enstaka undantag är svårare att försvara



Figur 1. Evidensbaserad medicin samspelar med – och stärker – andra viktiga komponenter i läkekonsten men kan inte ersätta dessa. Modifierat från Haynes och medarbetare [16].

men kan ändå vara någorlunda klokt, att ständigt göra undantag går inte att rättfärdiga.

I all sjukvård finns en osäkerhet. Denna osäkerhet måste vi lära oss att hantera. För sig själv och inför patienten måste man öppet kunna erkänna gränserna både för den egna kunskapen och för det medicinska vetandet, något som beskrivits som en grundsats i »ethics of evidence« [6].

I det praktiska handlandet blir balansgången ofta svår när det vetenskapliga underlaget är mycket svagt eller rentav obefintligt (se exempel 1). Många av oss som förespråkar EBM har försökt lösa dilemmat genom att tala om att beslut kring enskilda patienter måste grundas »på bästa tillgängliga vetenskapliga underlag« och att man måste vara medveten om att detta underlag ibland är mycket bräckligt [2, 5]. Men vi måste samtidigt inse att EBM saknar en klar strategi för hur man ska hantera dessa svåra kliniska situationer. Vanligen hänvisar man till att det krävs mer klinisk forskning på området. Detta är ett långsiktigt mål som är lätt att instämma i, men det hjälper inte den patient som behöver hjälp idag.

Hur hantera världens svårfångade värden?

En annan svårighet när man söker tillämpa EBM i praktisk klinisk verksamhet rör de svårfångade värdena i vården – det finns i utövandet av läkekonsten element som är svåra, ibland omöjliga, att kvantifiera eller uttrycka explicit [4]. Man kan förvisso komma långt i ansträngningarna att göra mjuka fakta hårda, det som kallats clinimetrics [7]. Det kan gälla analogskalor för smärta eller förtvivlan, poängsystem för neurologiska funktioner, ADL-förmåga och livskvalitet eller algoritmer och dataprogram för klinisk diagnostik utifrån anamnes och fysikalisk undersökning. Den kvalitativa forskningen har tillfört ytterligare dimensioner. Men också den kan handla bara om sådant som kan göras explicit, uttryckas i tal och skrift. Kvar finns kliniska fenomen och komponenter i samspillet mellan läkare och patient som inte fångas av kvantitativa instrument eller kvalitativa forskningsansatser [8]. Det är uppenbart att EBM ger lösningen på bara delar av de problem som läkaren möter i sitt dagliga arbete.

Kanske kan man gå till moralfilosofin för att finna en grund för distinktionen mellan klinisk praxis och den vetenskapliga grunden för vårt kliniska handlande. Här skiljer man nogsamt mellan fakta och värderingar. Tillsammans bestämmer de våra etiska handlingar [4]. Enbart fakta kan sällan eller aldrig bestämma hur vi ska handla i svåra beslutssituationer. Omvänt blir kliniska beslut svåra att fatta när fakta saknas – ju bättre faktaunderlag, desto lättare blir det för läkare och patient att avgöra hur man ska handla.

Detta gör att det egentligen inte finns någon motsättning mellan det vetenskapliga underlaget å ena sidan och läkarens och patientens erfarenheter och värderingar å den andra (Figur 1). För god läkekonst är kunskapsunderlaget oerhört vik-

II Exempel 1

På min mottagning på medicinkliniken möter jag ibland patienter med kroniskt trötthetssyndrom. Dessa patienter har inte sällan mycket svår funktionsnedsättning, klarar inte arbete vare sig utanför eller i hemmet och har ett torftigt socialt liv. Kognitiv beteendeterapi har i randomiserade prövningar visat sig vara av visst värde, även om den inte ger bot [14]. En del av de patienter som genomgått kognitiv terapi tycker sig inte hjälpta, andra vill inte ens pröva.

Vad återstår då? Bland patienterna finns en stor tilltro till megadoser av mekobalamin (vitamin B₁₂) som terapi. Några av mina patienter har fått denna behandling intramuskulärt 1–3 gånger per vecka. De flesta känner sig betydligt piggare och tycker sig ha fått förbättrad livskvalitet. Vid utsättningsförsök anger de klar försämring.

Varken jag eller patienterna vet om detta är en reell läkemedelseffekt eller om det rör sig om en kraftfull placeboeffekt. Det saknas helt underlag från randomiserade kliniska prövningar om effektiviteten hos denna terapi. Den bygger i stället på bräckliga observationer att homocysteinmetabolismen i centrala nervsystemet skulle vara störd hos patienter med kroniskt trötthetssyndrom [15]. En extensiv litteraturgenomgång har inte kunnat finna några kända biverkningar av höga mekobalamin-doser utöver akne.

Min behandling är ett klart avsteg från EBM-principerna. Situationen förvärras av att jag är styrelseordförande för SBU, ett av den evidensbaserade medicinens starkaste fästen i landet. Skulle jag vilja mer storskaligt utöva denna terapi måste jag väl göra det inom ramen för en randomiserad prövning? Men i mötet med den enskilda patienten – och så länge vi båda erkänner den osäkerhet som finns kring det vetenskapliga underlaget – är detta avsteg från EBM berättigat eller inte?

tigt – och varför ska inte det vara så bra som möjligt? Men i vår entusiasm för EBM får inte de andra komponenterna i läkekonsten skymmas. Det kan tyckas vara något av en paradox att det finns åtskilligt vetenskapligt underlag för vad som främjar eller motverkar dessa icke-faktabaserade delar av läkekonsten, något som SBUs stora genomgång över patient–läkarrelationen visat [9].

Nu är många av förespråkarna för EBM långtifrån omedvetna om gapet mellan klinisk-epidemiologisk forskning och klinisk praxis. I den kliniska forskningen ägnar man åtskillig uppmärksamhet åt värderingar, och då mest patientens och anhörigas värderingar, som grund för kliniska beslut. Detta gör man på sant EBM-manér, genom kvantifieringar. Den vanligaste formen är att i s k cost–utility-analyser använda sig av livskvalitetsjusterade levnadsår (QALYs). Dessa tekniker kan fungera på gruppnivå. Att tillämpa dem i klinisk praxis ter sig inte lika okomplicerat, något som exempel 2 visar.

EBM som underlag för prioriteringar

Svårigheterna kring hur man ska förfara när det vetenskapliga underlaget saknas eller är otillräckligt gäller inte bara den kliniska konsultationen, de gäller i minst lika hög grad prioriteringar på politisk-administrativ nivå.

Grovt kan sägas att det finns mycket god vetenskaplig dokumentation kring läkemedel, god eller halvgod dokumentation av kirurgiska insatser, svag vetenskaplig dokumentation av värdet av omvårdnadsinsatser samt inte bara svag utan också kontroversiell dokumentation av rehabiliteringsmetoder (att det finns enstaka lysande undantag från denna generalisering är uppenbart). Skevheten i det vetenskapliga under-

II Exempel 2

För att i vetenskapliga studier mäta det värde en försöksperson sätter på olika sjukdomstillstånd eller behandlingar, kan han/hon ställas inför ett hypotetiskt beslut:

»Föreställ dig att du har en svår sjukdom. Du står framför två dörrar. Bakom den ena väntar ett liv med invaliditet, du sitter i rullstol men kan äta själv och tala. Bakom den andra dörren väntar en farlig operation. Den ger dig 70 procents chans att bli helt återställd och leva ett fullständigt normalt liv, men också 30 procents risk att avlida vid operationen. Vilken dörr väljer du?»

Beroende på försökspersonens svar varierar sedan procenttalen tills han/hon inte längre kan välja säkert. På detta sätt når man ett »utility«-värde, vilket sedan kan användas i s k cost-utility-analyser.

Denna metod kan fungera bra som forskningsinstrument, men är knappast vad man skulle vilja tillämpa i praktisk sjukvård för att få en uppfattning om patientens värderingar.

laget inom sjukvårdens olika delar får konsekvenser för prioriteringsbeslut. Om en av grunderna för prioriteringar ska vara vetenskaplig dokumentation av metodens/åtgärdens värde, så hamnar onekligen läkemedlen i den gynnade klassen, omvårdnaden och rehabiliteringen i stryckklassen. Detta är nog en av anledningarna till att kritiken mot EBM ibland är särskilt engagerad från omvårdnadshåll [10]. Man framhåller att kvantitativa forskningsmetoder, ofta använda inom omvårdnaden, inte lämpar sig för den typ av systematiska kunskapsöversikter och metaanalyser som kännetecknar EBM.

Just denna kritik mot EBM är knappast berättigad. Också för kvalitativ forskning kan man ställa upp strikta kvalitetskrav, och det är fullt möjligt att göra systematiska kunskapsanalyser (men givetvis inte metaanalyser med sammanfattande statistiska mått) [11]. Det går att avgöra om det finns ett tillräckligt gott vetenskapligt underlag också inom områden som domineras av kvalitativa forskningsansatser.

Omvårdnadens och rehabiliteringens problem är i stället att det vetenskapliga underlaget är så svagt och att man ägnat så mycket kraft åt deskriptiv forskning, medan effekterna av olika interventioner mer sällan är utforskade. Ett exempel: Genom att tidsåtgången är så stor och metoden tillämpas så storskaligt är IVP (individuell vårdplanering) troligen den mest resurskrävande metod som införts i svensk sjukvård under de senaste decennierna. Det vetenskapliga underlaget är mycket bräckligt för att avgöra om metoden är till nytta eller inte, vilka bieffekter den kan ha, hur kostnadseffektiviteten är etc. Det finns en rad studier som kartlagt vårdpersonalens tillfredsställelse med metoden, men storskaliga prövningar av metoden och dess värde för patienterna saknas helt [12].

På motsvarande sätt finns stora svårigheter när det gäller att tillämpa EBM-principerna på organisatoriska beslut. En systematisk genomgång av det vetenskapliga underlaget för storskaliga organisatoriska förändringar inom sjukvården konstaterar att detta underlag är mycket svagt [13]. Artikeln slutar uppgivet: »These exercises continue to be repeated, perhaps as a tactic for creating an illusion of progress where none in reality exists.« För storskaliga organisatoriska förändringar kan man sällan hänvisa till det vetenskapliga underlaget, eftersom det saknas. Andra typer av argument används i stället.

För prioriteringar är EBM-principerna således till störst hjälp när det gäller att väga mellan olika insatser inom en och samma sfär inom sjukvården, t ex mellan olika läkemedel (Trombyl, Plavix eller Waran som trombosprofylax?) eller

olika kirurgiska ingrepp (ballongdilatation eller koronar by-pass?). De kan också vara av värde när det gäller beslut om organisatoriska insatser på basnivå (stroke-enhet? ryggskola? rökavvänjningsklinik?). EBM är, nästan definitionsmässigt, till föga hjälp när det gäller prioriteringar inom områden där det vetenskapliga underlaget är mycket svagt eller saknas, som inom stora delar av omvårdnad, rehabilitering och sjukvårdsorganisation.

Randomiserade prövningar ger inte alla svar

EBM litar i mycket stor utsträckning till resultat från storskaliga randomiserade prövningar. Denna metodik är oöverträffad när det gäller att fastställa om en diagnostisk eller terapeutisk insats är av värde för grupper av patienter eller inte. Konceptuellt rör sig dessa studier i gränslandet mellan klinisk medicin å ena sidan och epidemiologi å den andra – »klinisk epidemiologi« studerar större grupper av patienter (i motsats till enskilda individer).

Den kliniska medicinens centrala uppgift är att vara enskilda patienter till gagn – det handlar om en mycket personlig relation, hur en människa söker hjälp och stöd hos en annan. Epidemiologin/samhällsmedicinen/folkhälsovetenskapen har ett annat mål: att göra insatser som kommer hela befolkningen till godo. Den kliniska medicinens och folkhälsovetenskapens mål utesluter inte varandra, men de är klart åtskilda och kan ibland komma i konflikt med varandra.

Storskaliga kliniska prövningar och populationsbaserade epidemiologiska studier lär oss mycket om grupper av patienter och om befolkningar. Men de lär oss inte nödvändigtvis lika mycket om enskilda patienter. I de kliniska prövningarna söker man eliminera variationerna mellan individerna med randomisering, men det betyder självfallet inte att dessa variationer är betydelselösa. Det är inte bara effekterna av variationer mellan patienter man söker avlägsna genom randomiseringen. Kanske lika viktigt är att randomisering och dubbelblindförfarande minimerar effekterna av skillnader mellan läkare/terapeuter. Just de icke-kvantifierbara faktorer som randomiseringen söker eliminera kan ändå vara helt avgörande i det unika mötet mellan läkaren och patienten.

Förskjutningar i den kliniska verksamhetens fokus

Om man bortser från EBMs begränsningar, hamnar man i svårigheter. Det finns risk för två slag av tyngdpunktsförskjutningar:

- Sjukvården blir riktlinje- och algoritminriktad, styrd av fakta som gäller för grupper av patienter, till förfång för individens värde – patientens likaväl som doktors. Litar man uteslutande till resultatet från randomiserade prövningar bortser man från den betydelse de individuella variationerna har.
- Fokus förflyttas, på gott och ont, från det som är till gagn för den individuella patienten till samhällsnyttan.

Till det senare – samhällsnyttan – bidrar den komponent som ursprungligen inte var förknippad med EBM men som alltmer kommit att bli det, nämligen kostnadseffektivitetsaspekten. Den allmänt accepterade definitionen av EBM talar om »... bästa tillgängliga vetenskapliga underlag till gagn för enskilda patienter« [2]. Men när beräkningar av kostnader görs inom EBM-ramen har dessa i regel ett samhälleligt perspektiv. Visserligen kan den enskilda doktorn rättfärdiga detta perspektiv – själv brukar jag använda orden »som läkare är det min plikt att se till att resurserna räcker inte bara till den patient jag har framför mig utan också till nästa«. Men konflikten finns onekligen där mellan individ- och samhällsperspektivet i kostnadseffektivitetsanalyserna.

På sina håll, t ex i USA och Tyskland, inför man ekono-

miska incitament för den enskilda läkaren att begränsa kostnaderna, ofta med EBM som argument. Detta inkräktar förmodligen på grunderna för relationen mellan läkare och patient – det kan bli oklart om ett kliniskt beslut i första hand görs i läkarens eller i patientens intresse [6]. I den offentliga vården i Sverige anger de administrativa beslutsfattarna ramar för de resurser som finns tillgängliga för läkaren och patienten, vilka inom dessa ramar söker det som är bäst för patienten. Detta ingriper mindre i patient–läkarrelationen men förutsätter en stor lojalitet med de övergripande ekonomiska målen från läkarens sida för att systemet ska fungera.

Referenser

1. Stolt CM. Evidensbaserad medicin: Tio kritiska invändningar mot en värdefull metod. *Läkartidningen* 2001;98:3255-6.
2. Sackett DL, Richardson WS, Rosenberg, WM, Haynes RB. How to practice and teach evidence-based medicine. 2nd ed. New York: Churchill-Livingstone; 2000.
3. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implications. *Acad Med* 1990;65:611-21.
4. Tonelli MR. The philosophical limits of evidence-based medicine. *Acad Med* 1998;73:1234-40.
5. Rosenberg W, Donald A. Evidence-based medicine: an approach to clinical problem-solving. *BMJ* 1995;310:1122-6.
6. Miké V. Outcomes research and the quality of health care. The beacon of an ethics of evidence. *Eval Health Prof* 1999;22:3-32.
7. Sledge WH, Feinstein AR. A clinimetric approach to the components of the patient-physician relationship. *JAMA* 1997;278:2043-8.
8. Goldman GM. The tacit dimension of clinical judgement. *Yale J Biol Med* 1990;63:47-61.
9. Ottoson JO, red. Patient–läkarrelationen. Läkekonst på vetenskaplig grund. Stockholm: SBU och Natur & Kultur; 1999.
10. Clarke JB. Evidence-based practice: a retrograde step? The importance of pluralism in evidence generation for the practice of health care. *J Clin Nurs* 1999;8:89-94.
11. Popay J, Rogers A, Williams G. Rationale and standards for the systematic review of qualitative literature in health services research. *Qual Health Res* 1998;8:341-51.
12. Asplund K, Willman A, Alton V, Norberg A. Individuell vårdplanering – vad är det vetenskapliga underlaget? *Vård i Norden* 2000;20:46-9.
13. Ferlie E. Large-scale organizational and managerial change in health care: a review of the literature. *J Health Serv Res Policy* 1997;2:180-9.
14. Prins JB, Bleijenberg G, Bazelmans E, Elving LD, de Boo TM, Severens JL, et al. Cognitive behaviour therapy for chronic fatigue syndrome: a multicentre randomised controlled trial. *Lancet* 2001;357:841-7.
15. Regland B, Andersson M, Abrahamsson L, Bagby J, Dyrehag LE, Gottfries CG. Increased concentrations of homocysteine in the cerebrospinal fluid in patients with fibromyalgia and chronic fatigue syndrome. *Scand J Rheumatol* 1997;26:301-7.
16. Haynes RB, Sackett DL, Gray JM, Cook DJ, Guyatt GH. Transferring evidence from research into practice: 1. The role of clinical care research evidence in clinical decisions. *ACP J Club* 1996;125:A14-A16.

Särtryck

Läkartidningen

Nya vetenskapliga rön har lagt grunden till en ny klassifikation, nya diagnoskriterier, effektivare behandling och prevention – inklusive vaccination – av diabetes, som uppvisar en närmast epidemisk spridning. Målen och medlen för den snabba förbättringen av diabetesvården som nu är möjlig belyses i *Läkartidningens* serie, som också analyserar konsekvenserna för patienterna, sjukvården och samhället av de förändringar som Socialstyrelsen hösten 1999 fastställde som nationella riktlinjer.

Priset är 85 kronor.



Diabetes

Beställer härmed ex
av "Diabetes"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till Läkartidningen
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

kom från Landstingsdirektörernas förening som också varit uppdragsgivare för det fortsatta arbetet.

Med inspiration från denna resa lyckades Holmquist och Schützer entusiasmera Landstingsförbundet till ett projekt med syftet att skapa ett lättillgängligt digitalt bibliotek för all sjukvårdspersonal. I projektgruppen har, utöver idégivarna, också undertecknad haft förmånen att medverka tillsammans med Ida Noppa från Vårdförbundet och Eva Alopeus från Sahlgrenska universitetsbiblioteket. Arbetsgruppens ledare har varit Jan Lindmark på Landstingsförbundet som med stort tålamod och målmedvetenhet lotsat upphandlingen i hamn.

Processen har varit lång och vägen ofta törnbeströdd, men arbetet har nu nått målet att kunna introducera ett lättillgängligt digitalt bibliotek i stora delar av Sjukvårdssverige.

Eira – sjukvårdens digitala bibliotek

Ett 60-tal av de mest utnyttjade vetenskapliga tidskrifterna samt fem viktiga databaser finns redan från början tillgängliga på samtliga datorer i 13 landsting – Blekinge, Östergötland, Sörmland, Värmland, Västmanland, Kronoberg, Kalmar län, Jönköping, Uppsala, Dalarna, Jämtland, Västerbotten och Gävleborg – samt i Gotlands kommun. Tillgänglighet och enkelhet har varit viktiga lösenord i arbetet med att bygga upp systemet.

Det digitala biblioteket är tillgängligt för all personal och nås via en förhoppningsvis enkel och lättarbetad sökmotor; adressen är: <http://www.lf.se/eira/>. Det går enkelt att söka information om ett aktuellt kliniskt problem eller om kunskapsläget kring logistiska problem i sjukvården. Resultatet kan självfallet skrivas ut på papper om man inte vill läsa direkt på skärmen. En viktig strävan har varit att artikeln ska se ut som i den gamla klassiska papperstidningen.

För att påminna oss om vår historiska bakgrund bär det digitala biblioteket namnet Eira efter en gudinna i asatron. Hon var speciellt bra på läkekonst. Namnet är fornnordiskt och betyder »förskoning«.

Ökade krav på biblioteken

Lättillgänglig information och nya grupper som får möjlighet att söka information kommer att öka kraven på sjukhusbiblioteken. Från att tidigare ha förmedlat artikelkopior och arbetat med att söka vetenskapligt underlag på frågor från en ganska begränsad grupp i sjukvården kan bibliotekens roll nu komma att förändras.

Kraven på utbildning i informations-sökning och hjälp med att skaffa veten-



Det »digitala biblioteket« Eira nås på adressen: <http://www.lf.se/eira/>

skaplig dokumentation kommer att öka och behöva förmedlas till en förhoppningsvis snabbt växande grupp sjukvårdspersonal. På sikt kommer denna kunskap också att efterfrågas av patienter via sjukhusbiblioteken, vilket adderar en ny och spännande målgrupp. Förhoppningsvis kommer också våra politiker och beslutsfattare att söka sig till ett vetenskapligt baserat informationsunderlag som stöd i sitt arbete med utveckling och prioritering i sjukvården.

Varför är inte alla med?

Sveriges största sjukvårdshuvudmän saknas i denna första upphandling. Politiska skäl har utgjort ett hinder på ett sätt som vi inte sett tidigare. Den ökade självständighet som sjukvårdsområden och sjukhus har i t ex Stockholm och Region Skåne har försvårat samordning av en gemensam upphandling på riksplånet.

Samtidigt är ett enskilt landsting – om än stort – sannolikt en för liten enhet för att rationellt kunna driva igenom en upphandling av denna karaktär. Detta är ett exempel på en önskad bieffekt av nya system i en föränderlig verklighet.

Det är viktigt att de som ska använda och är i behov av kunskapsförmedlingen i ett digitalt bibliotek hjälper till att skapa förutsättningar för att detta ska kunna spridas även i regioner med nya styrformer för sjukvården.

I några landsting har också en motvilja från sjukhusbiblioteken mot denna nya publikationsform bidragit till att de står sig utanför.

Hur ska det användas?

När ska det digitala biblioteket användas? Från min horisont kan jag se tre olika huvudområden för användning:

- Den kliniska verkligheten – ett patientrelaterat problem uppstår; när som helst på dygnet går det att söka och ta fram senaste vetenskapliga underlag för att fatta ett riktigt beslut.
- Ett annat område där enkel och översiktlig tillgång till litteratur underlättar är utarbetandet av vårdprogram, kliniska

riktlinjer, justeringar av vårdkedja och sjukvårdsplanering.

- Sist men inte minst har vi vår egen förkovran och fortbildning där det digitala biblioteket kan ge enkel tillgång till lättlästa och trevliga artiklar.

Läkaren och kunskapen

Samhällsutvecklingen som helhet går mot allt bättre utbildade och krävande patienter. Även andra yrkesgrupper inom sjukvården skaffar sig fördjupade kunskaper inom sina respektive områden. Allt detta kommer att bidra till att läkarens »kunskapsprivilegium« blir ifrågasatt.

Det digitala biblioteket kommer att bidra till denna utveckling. Personligen är jag övertygad om att detta är bra och positivt. Genom att söka kunskap kan vi utvecklas och möta en spännande framtid som ställer ökade krav på vårt kunskapsutbud. Jag hoppas det digitala biblioteket ska bidra till detta. •

Evidensbaserad medicin – en god idé som hänger i luften

Evidensbaserad medicin (EBM) innebär att klinisk medicin skall bygga på bästa möjliga vetenskapliga underlag integrerat med den enskilda läkarens kliniska kunskap och erfarenhet. Det fordrar en organisation som erbjuder förutsättningar för eftertanke, utvärderingar, syntes av information, samtal och utbyten av erfarenheter. Den moderna sjukvårdsorganisationen saknar dessa förutsättningar. Att detta faktum inte med kraft diskuteras av begreppets ambassadörer leder till oro och hotar dess trovärdighet. Det är därför viktigt att EBMs förespråkare lämnar sina idealistiska positioner och diskuterar begreppet insatt i det sammanhang där de menar att det skall tillämpas.

ÅKE THÖRN

överläkare, med dr, yrkesmedicinska enheten, Sunderby sjukhus, Luleå
ake.thorn@nll.se

I nästan 200 år har två skilda idéer medverkat till både läkarkårens självbild och utvecklingen av modern medicinsk vetenskap och formerna för dess praktiska tillämpning. Den ena av idéerna har sina rötter i upplysningstidens ideal. Den andra är rotad i en romantiskt förankrad tradition som beskrivits som svärmisk, opraktisk och hätsk mot naturvetenskapliga nymodigheter.

Striden mellan upplysning och romantik

I en avhandling [1] om den på 1800-talet i Åbo och Uppsala verksamme medicinprofessorn Israel Hwasser har idéhistorikern Sven-Eric Liedman skildrat en strid mellan dessa två tanketraditioner. Den förstnämnda gick segrande ur konflikten och ledde till dagens vetenskapligt förankrade medicinska praktik. Men den andra förlorade trots allt inte helt utan kom att ligga till grund för den bild av läkaren som länge blev förhärskande: en central person, självskriven auktoritet för den sjuke och utrustad med ett stort samhälleligt anseende och inflytande.

Läkarrollen följer samhällsutvecklingen

Liedman framhåller hur, trots att ingen av de många läkare som skrivit om Hwasser instämt i dennes medicinska teorier, ingen heller »undvikit att slösa beröm över hans läkaretik, hans kamp för läkarens ställning, kort sagt hans läkarideologi« [2]. Sedan Hwassers tid har emellertid läkarens roll förändrats kontinuerligt. Den faderligt auktoritativa doktorsrollen passade så småningom inte längre ihop med vare sig den mo-

dena, rationella driften av sjukvården eller samhällsutvecklingen i övrigt.

Ett yrke i ständig kris?

Motzi Eklöf [3] vid Linköpings universitet har i en doktorsavhandling om den svenska läkarkårens identiteter, intressen och ideal visat hur läkaryrket från slutet av 1800-talet till 1960 ständigt tycks ha varit i en kris. Denna har handlat om förändringar av yrkets självbild och anseende, och om försök att samtidigt både tillvarata den traditionella synen på yrkesrollen och möta nya tiders krav på medicinsk behandling och vård.

Dessa försök att förena tradition och förändring har emellertid inte lyckats. Den gamla Hwasserska doktorsrollen har gradvis tonat bort och förändrats i riktning mot rollen som specialiserad tekniker med uppgift att reparera trasiga organ och/eller utfärda regler för hälsans bevarande. År 1975 förutspådde Liedman att denna utveckling småningom skulle gå ända dithän att läkaren bleve en »helt opersonlig dator, som tickar fram små remsor av diagnoser och recept« [2].

Vantrivsel med läkarrollen

Motzi Eklöfs studie går fram till 1960. Det är dock högst troligt att krisen fortsatt också därefter och kanske till och med förvärrats. Åtminstone har det under 1900-talets avslutande decennium bland läkare märkts en tilltagande vantrivsel med den allt snävare yrkesrollen och med organisationen som drivit fram den.

Det finns många tecken på detta. Sjukskrivningar bland läkare har ökat [4]. Massmediala larmrapporter om stressrelaterade sjukdomar i yrkesgruppen har blivit vanliga, och många läkare uppger sig vilja lämna yrket [5]. Arbets-

Evidensbaserad medicin

miljöinspektionen har bedömt arbetsmiljön som så pressad att många läkare balanserar på marginalen till att få arbetsskador [6]. Nyligen har en stor studie visat att läkare är överrepresenterade bland de anställda inom vården som haft sämst hälsoutveckling [7].

Förväntningarna har inte infriats

Yngre läkare har i Läkartidningen vittnat om besvikelse över att yrket inte varit det man trott, att organisationen inte tillåter reflektion och kreativitet utan snarast befrämjar en själlös löpbandproduktion [8]. Liknande vittnesbörd har också förts fram av deras äldre, mer etablerade kolleger som till exempel ordföranden i Svenska Läkaresällskapet, professor Birgitta Strandvik. Hon framhåller att utvecklingen lett till för mycket av industriellt tänkande och industriella produktionsmått i vården. Den stora brist på tid till reflektion och syntes av information som detta fört med sig bedömer hon nu hota kvaliteten inom såväl sjukvård som klinisk forskning [9].

Vetenskap eller läkekonst?

Eklöf visar att diskussioner om läkarens självbild, identitet och kris ständigt förts inom läkarkåren kring några få återkommande teman. Avsikten har varit att finna en doktorsroll som, med nödvändig hänsyn till tradition och position, bättre skulle passa nya tiders krav på läkaren och därmed lösa krisen. Ett av de återkommande ämnena har varit huruvida

Det är på tiden att EBMs förespråkare lämnar sina ideala positioner och diskuterar begreppet insatt i de sammanhang där det skall användas!

tyngdpunkten i läkarens verksamhet ska ligga på vetenskapliga fakta eller på en mer intuitiv läkekonst. Just detta är för närvarande ett dominerande tema, men inte som ett antingen–eller utan som ett både–och.

I den diskussion som under senare år offentligt förts bland läkare om yrkets kris har det således markerats ett starkt behov av en renässans för läkekonst som en väsentlig del i yrkesrollen [9, 10]. Samtidigt har frågan om den medicinska vetenskapens roll i den kliniska praktiken också bearbetats och diskuterats. Den har i hög grad kopplats till begreppet evidensbaserad medicin (EBM) som grovt definierats som »en strävan efter att integrera den enskilda läkarens kliniska kunskap och erfarenhet med bästa tillgängliga vetenskapliga underlag från externa källor« [11].

Formulerad på det sättet bör definitionen rimligen tolkas som att EBM handlar om en sammansmältning av läkekonst och vetenskap. Genom att stimulera till arbetssätt med en bättre balans mellan läkekonst och läkarvetenskap skulle det i så fall kanske kunna förlösa konstruktiva idéer om vägar ut ur läkaryrkets och sjukvårdens allvarliga kris [12].

Oro att evidensbaserad medicin tar över

Medan veterligen ingen under de senaste årens debatt öppet uttryckt farhågor för problem förknippade med en läkekonstens renässans så har många markerat oro inför EBMs pågående införande i sjukvården. I grunden verkar det gälla farhågor över att EBM helt skall komma att normera, styra och diktera medicinsk behandling och vård ända ner till det enskilda fallet [13–15].

Rädslan är olycklig därför att den kan förhindra det konstruktiva användandet av de synsätt som ligger bakom begreppet [12]. Den är dock begriplig och botten förmodligen i oklarheter kring definitionen som inte alls är så entydig som ibland görs gällande [16]. I SBU:s ordlista på Internet definieras t ex EBM enbart som en »medveten och systematisk användning av bästa tillgängliga vetenskapliga fakta«, och inte dessutom som en integration av detta med »den enskilda läkarens kliniska kunskap och erfarenhet« [11].

EBMs förespråkare, t ex SBU:s chef Mona Britton [17], rangordnar också onödigt den randomiserade kliniska prövningen som den främsta bland alla vetenskapliga metoder. Rangordningen görs oftast utan att man samtidigt påtalar metodens begränsningar, vilka bl a utgörs av de stora problemen med att generalisera standardiserad kunskap till att vara tillämplig på enskilda fall.

Intrycket blir därför lätt att EBMs

Läkare undersöker patient. Illustration i »Liber notabilium Philippi Septimii«, av Guy de Pavia, år 1345. Musee Conde, Chantilly, Frankrike.

talesmän tillskriver metodens resultat en universell överlägsenhet oberoende av klinisk situation, och att de anser den enskilda läkarens kliniska erfarenhet vara något mycket underordnat.

Mer nyanserad innebörd

Samtidigt kan oron tyckas obefogad och svår att begripa därför att andra framträdande språkrör, också knutna till SBU, upprepade gånger givit begreppet en mer nyanserad innebörd. Lars Werkö har till exempel gjort det indirekt genom en artikel där han skildrat hur man konkret kan arbeta med EBM i klinisk verksamhet. I den framskyntar visionen om en systematisk, kritiskt reflekterande seminarieverksamhet som en ständigt pågående del av den kliniska rutinen [18]. Han har också direkt, utan rangordning av vetenskapliga metoder, klargjort att EBM helt enkelt innebär (det självklara) att man skall veta på vilka grunder man fattar beslut inom sjukvården [19].

Mats Eliasson, SBU, har genom valet av de många skiftande perspektiven och infallsvinklarna i artiklarna i den serie om EBM i Läkartidningen som han redigerar, visat att begreppets grunder är många och går långt bortom den randomiserade kliniska prövningen. Han har också i svar till skribenter med farhågor klart deklarerat att EBM endast är ett hjälpmedel för klinisk verksamhet och att andra färdigheter och kunskaper också krävs [20].

Man kan alltså fråga sig varför det ändå finns en så stark och återkomman-

Foto: IBL Bildbyrå

de oro över att EBM skulle normera medicinsk behandling till att vara strikt mekanisk, baserad på de kliniska randomiserade prövningarnas generella genomsnittskunskap. Varför läggs inte den energin i stället på att till exempel diskutera, utveckla och pröva former för systematiska, reflekterande kunskaps- och erfarenhetsutbyten kring kliniska frågeställningar i rutinsjukvården?

Idealistiska förespråkare

Mer än på den ibland otydliga definitionen beror detta sannolikt på att EBM-förespråkarna är märkligt idealistiska. Trots deras starka intresse och bekännelse till vetenskaplig evidens sätter de dessvärre inte in, diskuterar och analyserar begreppet i det sammanhang där det skall tillämpas, nämligen i den sjukvårdsorganisation som det senaste decenniet präglats av ständiga, snabba förändringar. Dessa har i största utsträckning dominerats av teoretiska perspektiv och konsultburna modeväxlingar med få kopplingar till empirisk forskning [7], och de har aldrig varit evidensbaserade.

De har sällan följts upp och när så skett har grova surrogatmått som behandling per vårdplats, per anställd eller per skattekrona använts medan patientrelaterade effektmått saknats.

Stress och sönderslagna nätverk

Detta betyder inte att förändringarna saknat effekter. De har medfört en kraftigt ökad arbetsbelastning, stress och sjukskrivning, upplevelse av otillräcklighet och skuldkänslor, konflikter och revirstrider och slagit sönder fungerande nätverk [7]. Initiativ och kreativitet har minskat och möjligen försvunnit [7, 8]. Tid har blivit sjukvårdsorganisationens största bristvara [9].

Det verkar med andra ord som om den reflekterande organisation som skulle kunna utveckla klinisk praktik baserad på en EBM som sammansmälter läkekonst med vetenskap egentligen inte finns i sinnevärlden. Detta minskar trovärdigheten för begreppets förverkligande i den breda bemärkelse det ges av Lars Werkö [19], Mats Eliasson [20] och andra med anknytning till SBU [11].

Det gör det också i hög grad begripligt att många läkare känner oro för att den sjukvårdsorganisation som faktiskt finns och dagligen upplevs, genom sitt funktionssätt kommer att tolka och praktiskt förverkliga EBM på snävast tänkbara sätt så att EBM blir den begränsande norm ingen önskar.

Läkarens oro och tveksamhet inför begreppet gäller därför antagligen farhågor om att det kan leda till att de slutgiltigt tvingas iklä sig den läkarroll som Liedman förutspådde år 1975: att de blir en

helt personlig dator som tickar fram små remsor av diagnoser och recept [2].

För att hindra en sådan utgång är det på tiden att EBMs förespråkare lämnar sina ideala positioner och diskuterar be- greppet insatt i de sammanhang där det skall användas!

Referenser

- Liedman SE. Israel Hwasser. Stockholm: Almquist & Wiksell, 1971.
- Liedman SE. Det vita skrået. Några notiser. Ord&Bild 1975;84:313-7.
- Eklöf M. Läkarens ethos. Studier i den svenska läkarkårens identiteter, intressen och ideal 1890–1960 [dissertation]. Linköping: Linköpings universitet, 2000.
- Filipsson A. Sjukdagar ökade med 40 procent – distriktsläkarna i topp i statistiken. Läkartidningen 2000;97:4272.
- Wilhelmson B. Var tionde läkare vill lämna vården. Ny kartläggning visar att läkares arbetsmiljö försämrats drastiskt. Läkartidningen 1998;95:3274-5.
- Rönngren M. Läkarna räcker inte till. Arbetskydd 1999;nr1:11.
- Petterson IL, red. Vårdens arbetshälso- rapport. Vårdarbetets villkor och konsekvenser för arbetshälsan. Stockholm: Stockholms läns landsting: Yrkesmedicinska enheten, 2000.
- Ahlgren T. Eva Nilsson Bågenholm, ny ord- förande i SYLF om läkaryrket: Kreativitet och frihet saknas. Läkartidningen 1998;95: 2972-3.
- Wilhelmson B. Birgitta Strandvik efterlyser paradigmskifte. När Mammon får styra ... vem tar ansvar för framtiden? Läkartidning- en 2001;98:3568-72.
- Johnson K. Läkare ska forma framtiden. Lä- kartidningen 1999;96:5477.
- Nordin-Johansson A, Asplund K. Evidens- baserad sjukvård i praktiken. Vetenskapliga grunden för kliniskt handlande bättre än sitt rykte. Läkartidningen 1999;96:3173-6.
- Mägi M. Är bara allmänläkarna vilna? All- mänläkarnas identitetskris – hela sjukvård- ens kris? Läkartidningen 2001;98:2228-9.
- Dalén P. Evidensbaserad medicin i vardags- arbetet. Överstepräster som visar oss den rätta, evidensbaserade vägen? Läkartid- ningen 2000;497-8.
- Nilsson G. Varning för alltför ensidig evi- densbaserad medicin! God läkarkonst för- utsätter även människokänedom och intui- tion. Läkartidningen 2001;98:2437-8.
- Stolt CM. Tio kritiska invändningar mot en värdefull metod. Läkartidningen 2001;98: 3255-6.
- Gårdelöf B. För femtielfte gången – evi- densbasering är inte detsamma som rando- miserade studier. Läkartidningen 2001; 98: 3164.
- Britton M. Så graderas en studies veten- skapliga bevisvärde och slutsatsernas styr- ka. Läkartidningen 2000;97:4414-5.
- Werkö L. Ett konkret redskap för vårdens verkstadsgolv. Läkartidningen 2000;97: 2710-3.
- Werkö L. Många kritiker missförstår EBM. Läkartidningen 2001;98:3449-50.
- Eliasson M. Välkommen debatt om evi- densbaserad medicin. Läkartidningen 1998;95 3450.

Evidensbaserad medicin, javisst, men var finns den?

Modern sjukvård ska baseras på bästa tillgängliga vetenskapliga underlag. En förutsättning är att sjukvårdspersonalen har tillgång till detta vetenskapliga underlag. I artikeln presenteras Eira – det nya lättillgängliga digitala bibliotek som introducerats i 13 landsting och i Gotlands kommun.

ULF GUNNARSSON
docent, avdelningsläkare, Akademiska sjukhuset, Uppsala
ulf.gunnarsson@kirurgi.uu.se

II Vetenskapliga tidskrifter och databaser utgör sedan länge det underlag som vi baserar vårt kunnande på. En viktig milstolpe i att göra detta kunskapsmater- ial tillgängligt var introduktionen av da- tabaserade sökningar genom MedLine för några decennier sedan.

Trots att antalet tidskrifter och artik- lar inom varje ämnesområde ökat lavin- artat under senare år har ändå möjlighete- rna att överblicka kunskapsstoffet för- bättrats. En betydande flaskhals har dock varit möjligheten att komma i be- sittning av kunskapen/dokumentatio- nen. Att be sjukhusbiblioteket skaffa fram en artikel och vänta kanske några veckor tills den anländer tar alltför lång tid för att vara rationellt i en effektiv sjukvårdsapparat.

Kunskap – en färskvara

Vårdprogram har traditionellt varit ett sätt att öka tillgängligheten till bearbetad kunskap. Svagheten hos dessa vårdprog- ram är att de är en färskvara. Redan efter något år kan rekommendationerna vara inaktuella. För att kunna omvärdera kun- skapsläget behövs enkel tillgång till det faktamaterial som tillkommit inom om- rådet. Dessutom tenderar det att uppstå situationer i den kliniska vardagen som inte omfattas av vårdprogrammen.

En snabb uppdatering är nödvändig för att kunna göra en korrekt värdering av bästa behandlingsstrategi. Tillgäng- ligheten till de studier och artiklar samt databasmaterial som behandlar den ak- tuella problematiken är avgörande för att snabbt kunna fatta riktiga beslut. När se- dan vårdprogrammen ska omarbetas el- ler en vårdkedja justeras är översiktlig och lätthanterlig tillgång till vetenska- pens senaste framsteg en förutsättning.

Lång dags färd mot ljus

På många amerikanska sjukhus har till- gången till söksystem, databaser och ve-

Evidensbaserad medicin

tenskapliga artiklar i den egna datorn länge varit en självklarhet. Sjukvård- enheter som konkurrerar med sitt rykte använder informationstillgängligheten som ett led i sin kvalitetsförbättring.

Läkare och forskare vid svenska uni- versitet har sedan flera år tillgång till ett välfungerande digitalt bibliotek. Från datorn på arbetsplatsen eller i hemmet kan de läsa och skriva ut artiklar från de flesta viktiga vetenskapliga tidskrifter. Deras kolleger i helt landstingsbaserad verksamhet har däremot nästan helt sak- nat denna möjlighet.

Initiativ för några år sedan

Växande kunskap skapar ökande behov. Flera kolleger har uppmärksammat den bristande tillgängligheten till snabb in- formation. Sjukhusbiblioteken har en nyckelfunktion i informationsförmed- lingen inom sjukvården. Även här har problemet uppmärksamats. För några år sedan reste två framsynta biblioteks- chefer, Bengt Holmquist i Falun och Jan Schützer i Karlstad, på en studieresa till bl a Mayo-klikerna i USA. Uppdraget

Processen har varit lång och vägen ofta törnbe- strödd, men arbetet har nu nått målet att kunna intro- ducera ett lättillgängligt digitalt bibliotek i stora delar av Sjukvårdssverige.

Immunglobulinbehandling vid svår sepsis och septisk chock saknar dokumenterad effekt

Cochranes översikt ett stolpskott!

■ En evidensbaserad Cochraneöversikt om behandling av sepsis och septisk chock med immunglobuliner utmynnar i en konklusion, där författarna anser att högdosbehandling med polyklonalt intravenöst immunglobulin (IVIG) signifikant visats reducera mortaliteten och därför – utan närmare specifikation – kan användas som adjuvant behandling vid dessa tillstånd [1]. Att översikten publicerats som en Cochrane-rapport tycker man borde borga för kvalitet i slutsatserna. Den aktuella rapporten är dock ett rörigt dokument, som dessutom vittnar om bristande insikt i den varierande patofysiologi som föreligger vid olika septiska tillstånd.

Vi kan lämna därhän författarnas mycket tvivelaktiga metod att slå samman polyklonala antikroppar både med monoklonala antikroppar riktade mot endotoxiner/cytokiner och med fysiologiska cytokinantagonister, som helt saknar immunglobulinstruktur, och kalla alltsammans för immunglobulinbehandling. Låt oss i stället koncentrera oss enbart på högdosbehandling med polyklonalt IVIG, där författarnas slutsats efter att ha granskat elva randomiserade kontrollerade studier är att behandlingen avsevärt reducerar dödlighet i sepsis/septisk chock. I dessa studier inkluderades totalt 492 patienter, men inte i någon av dem förekom mer än 30 behandlade patienter. Några av studierna var av tvivelaktig kvalitet, och när författarna bara tog med dem som ansågs vara av god metodologisk kvalitet baserades slutsatsen på endast 332 patienter. Cirka hälften av patientunderlaget utgjordes av barn med neonatal sepsis (Chen 1996, Shenoj 1999, Weisman 1992, Erdem 1993, Haque 1988; Figur 1) och återstående del huvudsakligen av vuxna patienter med postoperativ sepsis (De Simone 1988, Dominioni 1991, Grundmann 1988, Just 1986, Schedel 1991, Wesoly 1990; Figur 1).

Cochranes slutsats på skakig grund

Ingen av de ingående studierna var dimensionerad för att kunna besvara frågeställningen huruvida behandlingen hade effekt på överlevnad. Sepsis är inte något enhetligt tillstånd, och mycket olikartade inklusionskriterier har använts i de olika studierna. Detta är troligen huvudorsaken till att placebomortaliteten varierar så påtagligt – från 4 procent till 86 procent. Att författarnas heterogenitetstest ej fallit ut beror med största sannolikhet på de mycket låga patientantalen i varje studie. Genom att lägga ihop samtliga studier har författarna fått fram en signifikant behandlingseffekt för polyklonalt IVIG.

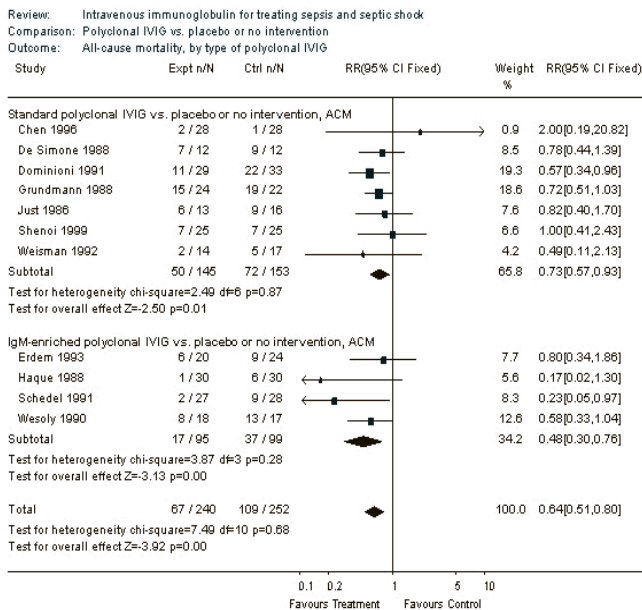
Evidensbaserad medicin

Om vi nu exempelvis fokuserar på vuxna patienter, där behandlingseffekten i rapporten redovisas som större än i neonatalgruppen, består det underliggande patientmaterialet av sex studier (De Simone 1988, Dominioni 1991, Grundmann 1988, Just 1986, Schedel 1991, Wesoly 1990). Två av dessa redovisar obalans i prognostiska faktorer till behandlingsgruppens fördel (Grundmann, Wesoly) och en studie har en anmärkningsvärt hög placebomortalitet i relation till inklusionskriterier och riskscore vid behandlingsstart (Dominioni). I en av studierna (Schedel) har man undersökt effekten av en IgM-berikad immunglobulinpreparation med hög halt av endotoxinantikroppar på patienter som före inklusion uppvisar ett positivt endotoxintest. Till de återstående studierna har så få patienter rekryterats att en tendens till positivt utfall för behandlingsgruppen avgörs av bara en eller två patienter (Just, De Simone). Det bör också poängteras att randomiseringsprocessen i fyra av studierna inte var blindad (De Simone, Just, Schedel, Wesoly). Sammanfattningsvis är således vår bedömning att författarnas slutsats vilar på mycket skakig grund.

Den största studien är inte med

Det kanske allvarligaste misstaget i Cochrane-rapporten är dock att man inte har fått med resultaten från den största och mest välgjorda studien i ämnet, den av Karl Werdan och medarbetare, vilken ännu inte publicerats. Prövningen startade 1991; de viktigaste resultaten redovisades redan 1997 [2]. Studien hade då slutförts, och någon positiv effekt av immunglobulinbehandling på överlevnaden hade inte påvisats. Sammanlagt hade 653 vuxna patienter med svår sepsis eller septisk chock inkluderats i denna stora studie. Om resultaten från denna hade tagits med i Cochranesammanställningen skulle slutsatsen troligen ha blivit helt annorlunda, dvs ingen effekt av polyklonalt IVIG vid svår sepsis/septisk chock hade påvisats.

Författarna till Cochrane-rapporten kan knappast ha varit okunniga om studien eftersom studieprotokollet till och med



Cochranesammanfattning av kliniska studier av IVIG vid sepsis/septisk chock.

Ur: Alejandria MM, Lansang MA, Dans LF, Mantaring JBV. Intravenous immunoglobulin for treating sepsis and septic shock [Cochrane Review]. In: The Cochrane Library, Issue 3, 2001. Oxford: Update Software [1].

hade publicerats [3], men på frågan om varför studien inte inkluderats i översikten uppger författarna som skäl att den ännu ej publicerats [1]. Denna Cochranerapport illustrerar således hur kraftigt publikationsbias kan slå mot en metaanalys. Små studier med garanterat dålig styrka och med hög känslighet för obalans i prognostiska faktorer men med tendens till positiv effekt blir publicerade, medan publiceringen av studier – även stora sådana – med negativt utfall dröjer eller kanske till och med inte alls blir av.

Publikationsbias ett universellt problem

Parallellen till studierna av profylax mot stressulcus och magblödning hos intensivvårdspatienter är slående:

En metaanalys av ett flertal mindre studier visade att sukralfat och H₂-blockerare båda minskade risken för ulcusblödning, men att sukralfatprofylax innebar mindre risk för respiratorassocierad pneumoni än användning av H₂-blockerare [4, 5]. När sedan den stora bekräftande studien med 1 200 inkluderade patienter kom till stånd kunde inte någon fördel med sukralfat påvisas [6]. Tvärtom gav H₂-blockerare en signifikant bättre effekt som ulcusblödningsprofylax, medan pneumoniriskerna var likvärdiga för de båda preparaten. Slutsatsen från metaanalysen baserad på ett flertal små studier kunde således inte bekräftas i den stora studien.

I metaanalysens barndom på 1980-talet missionerade bl a Richard Peto för att en grundlig översikt kräver att författaren genom sitt kontaktnät spårar upp även opublicerade randomiserade studier [7, 8]. Om så skulle behövas skall författaren själv till och med uppsöka enskilda forskare och sammanställa rådata från studier som blivit liggande av olika skäl. Den vanligaste orsaken till att kliniska prövningar inte publiceras är säkert att resultaten varit negativa både ur vetenskaplig och ibland måhända även ur kommersiell synvinkel.

Numera förefaller det som om vissa författare till metaanalyser ibland kan tycka att det räcker med en Medline-sökning och ett antal timmar framför datorn för att destillera fram san-

ningen. Även om sådana sammanställningar givetvis kan ha ett visst värde är svagheten påtaglig så länge problemet med publikationsbias kvarstår. En konstruktiv lösning som diskuterats, men tyvärr ännu inte genomförd, är dels att alla kliniska prövningar som startas måste registreras centralt, dels att det i villkoren för ett godkännande från etisk kommitté ingår att resultaten måste publiceras, alternativt efter en viss tid ställas till forskarsamhällets förfogande.

Polyklonalt IVIG utan visad effekt vid oselektad sepsis

Sammanfattningsvis anser vi – i motsats till konklusionen i Cochranerapporten – att polyklonalt IVIG ej tillfredsställande visats ha någon dokumenterad effekt i en oselektad patientpopulation med svår sepsis eller septisk chock, och att IVIG tills vidare därför ej heller bör användas på denna indikation. Svagheterna i konklusionen borde ha lyfts fram, och man borde i såväl sammanfattningen som i diskussionen ha informerat om existensen av Werdans studie och de preliminära resultaten därifrån. Tanken inställer sig onekligen att det kanske hade varit lika bra eller till och med bättre att helt avstå från denna metaanalys.

Avslutningsvis vill vi dock poängtera att ovanstående diskussion givetvis inte utesluter att polyklonalt IVIG kan ha effekt i vissa septiska subpopulationer. Exempel på en sådan är patienter med allvarlig grupp A streptokocksepsis med låga toxinneutraliserande antikropps nivåer [9]. Neutraliserande antikroppar finns i flera immunoglobulinberedningar, och kliniska studier har visat lovande resultat [10]. Den kliniska effekten har dock ännu ej slutgiltigt visats i en randomiserad kontrollerad prövning.

Referenser

- Alejandria MM, Lansang MA, Dans LF, Mantaring JBV. Intravenous immunoglobulin for treating sepsis and septic shock [Cochrane Review]. In: The Cochrane Library, Issue 3, 2001. Oxford: Update Software.
- Werdan K, Piltz G and the SBITS study group. Polyvalent immune globulins [abstract]. Shock 1997;7 Suppl 5.
- Piltz G, Fateh-Moghadam S, Viell B, Bujdoso O, Döring G, Marget W, et al. Supplemental immunoglobulin therapy in sepsis and septic shock – comparison of mortality under treatment with polyvalent iv immunoglobulin versus placebo. Protocol of a multicenter, randomized, prospective, double-blind study. Theoretical Surgery 1993;8:61-83.
- Cook DJ, Witt LG, Cook RJ, Guyatt GH. Stress ulcer prophylaxis in the critically ill: a meta-analysis. Am J Med 1991;91:519-27.
- Cook DJ, Laine GH, Guyatt GH, Raffin TA. Nosocomial pneumonia and the role of gastric pH. A meta-analysis. Chest 1991;100:7-13.
- Cook DJ, Guyatt GH, Marshall J, Leasa D, Fuller H, Hall R, et al. A comparison of sucralfate and ranitidine for the prevention of upper gastrointestinal bleeding in patients requiring mechanical ventilation. N Engl J Med 1998;338:791-7.
- Peto R, Halperin M, Brown B, Ellenberg S, Light R, Chalmers T, et al. Discussion. Stat Med 1987;6:229-31.
- Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. Journal of the Royal Statistical Society, series A 1988;151:419-63.
- Holm SE, Norrby A, Bergholm AM, Norgren M. Aspects of pathogenesis of serious group A streptococcal infections in Sweden. J Infect Dis 1992;166:31-7.
- Kaul R, McGeer A, Norrby-Teglund A, Kotb M, Schwartz B, O'Rourke K, et al. Intravenous immunoglobulin therapy for streptococcal toxic shock syndrome – a comparative observational study. Clin Infect Dis 1999;28:800-7.

Kommentar:

Kritikken er berettiget

II Bengt Gårdlund og Jan Sjölin er kommet med en vigtig kommentar til en Cochraneoversigt om en potentielt livreddende, men meget dyr behandling; behandling af sepsis og septisk chok med immunoglobuliner.

Vi er enige i, at Alejandria et al [1] burde have kontaktet Werdan et al [2] for at få adgang til data fra det store, upublicerede materiale, så det kunne indgå i overvejelserne. Og i, at det videnskabetiske komitéssystem bør kræve, at kliniske afprøvninger skal publiceres eller i det mindste stilles til rådighed for forskerne og de patienter, uden hvis velvilje det ikke er muligt at gennemføre kliniske forsøg [3, 4]. Vi støtter også det synspunkt, at Alejandria et al ikke har lagt tilstrækkelig vægt på, om »concealment of allocation«, blinding af randomiseringsprocessen, har været opretholdt.

Endelig er vi enige med Gårdlund og Sjölin i, at rimeligheden i at slå resultater fra specifikke og uspecifikke immunoglobuliner sammen kan diskuteres. Forfatterne har dog også anført resultaterne hver for sig.

Derimod mener vi ikke ...

Derimod mener vi ikke, at man principielt skal afstå fra at lave systematiske oversigter, hvis der kun findes nogle få, små studier af varierende videnskabelig kvalitet, der har beskæftiget sig med det spørgsmål, man er interesseret i. Når der skal træffes beslutning om, hvorvidt en behandling skal indføres eller ej, så vil beslutningstagerne alligevel forsøge at danne sig et overblik over den mest pålidelige dokumentation, der står til rådighed. Og hvis ikke der foreligger en systematisk oversigt, så vil alternativet ofte være en traditionel oversigt eller ekspertudtalelser med alle de risici for bias, dette indebærer [5].

Om det så indenfor rammerne af den systematiske oversigt er rimeligt at lave en metaanalyse eller ej, er en anden sag. Og naturligvis skal en systematisk oversigt baseret på små studier af tvivlsom kvalitet tolkes med stor forsigtighed og ydmyghed, hvilket desværre ikke altid er tilfældet, heller ikke for Cochraneoversigter [6].

En organisation, der skal vokse med opgaven

Dette eksempel på en Cochraneoversigt illustrerer, at selv om Cochranesamarbejdet gør meget for at levere de mest gennemskuelige, grundige, og videnskabeligt valide oversigter [7], så er det stadig en organisation, der skal vokse med opgaven [6]. Cochraneoversigter er vist at være af højere kvalitet og mindre behæftet med bias end systematiske oversigter publiceret i papirtidsskrifter [7-11], men det er helt afgørende, at gode kritikere som Gårdlund og Sjölin fremkommer med deres synspunkter, når det er påkrævet, således at oversigterne kan forbedres.

Vi vil opfordre Gårdlund og Sjölin til at indsende en kommentar til Cochraneoversigten (ved at trykke på Comments/Criticisms-knappen i CD-ROM-udgaven, eller på Feedback-knappen i Internet-udgaven af Cochrane Library), således at forfatterne får mulighed for at forbedre deres oversigt. Kommentaren vil da, efter redaktionel behandling, blive publiceret sammen med oversigten og forfatterens svar.

I dette tilfælde synes der dog at være brug for en mere omfattende og grundlæggende revision af oversigten. En af os

har tidligere indsendt en kritik af denne oversigt (se http://www.update.co.uk/feedback/comcrit_web.exe?S=clib.xml, og gå ned til *INFECTN*), og vi er på det Nordiske Cochrane Center i gang med at udarbejde en uafhængig systematisk oversigt om immunoglobuliner til behandling af sepsis og septisk chok.

Julie Pildal, læge, forsker

Peter C Gøtzsche, overlæge, dr med, direktør,

Det Nordiske Cochrane Center, København

Referencer

1. Alejandria MM, Lansang MA, Dans LF, Mantaring JBV. Intravenous immunoglobulin for treating sepsis and septic shock. Cochrane Review. In: The Cochrane Library 2001, issue 3. Oxford: Update Software.
2. Werdan K, Pilz Gat SSG. Polyvalent immune globulins [abstract]. Shock 1997;7:suppl 5.
3. Chalmers I. Underreporting research is scientific misconduct. JAMA 1990;263:1405-8.
4. Eliasson M, Bergquist D. Forskningsresultat bör vara allmänt tillgängliga! Läkartidningen 2001;98:3913-6.
5. Guyatt GH, Haynes RB, Jaeschke RZ, Cook DJ, Green L, Naylor CD, et al. Users' guides to the medical literature: XXV. Evidence-based medicine: principles for applying the users' guides to patient care. Evidence-Based Medicine Working Group. JAMA 2000; 284(10):1290-6.
6. Olsen O, Middleton P, Ezzo J, Gøtzsche PC, Hadhazy V, Herxheimer A, et al. Quality of Cochrane reviews: assessment of a sample from 1998. BMJ 2001;323:829-32.
7. Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M. Methodology and reports of systematic reviews and meta-analyses. JAMA 1998;280:278-80.
8. Shea B, Moher D, Pham B, Tugwell P. Assessing the quality of reporting meta-analyses of randomized controlled trials [abstract]. VII Cochrane Colloquium, Rome, 5-9 October 1999:A39.
9. Shea B, Dubé C, Moher D. Assessing the quality of reports of systematic reviews and meta-analyses: A systematic review of checklists and scales [abstract]. VII Cochrane Colloquium, Rome, 5-9 October 1999:A40.
10. Jadad AR, Moher M, Browman GP, Booker L, Sigouin C, Fuentes M, et al. Systematic reviews and meta-analyses on treatment of asthma: critical evaluation. BMJ 2000;320:537-40.
11. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997;315:629-34.

Placebo – något att tro på?

II Den amerikanske placeboforskaren Shapiro hävdar att den medicinska behandlingens historia fram till våra dagar i allt väsentligt har handlat om placeboeffekten [1]. Spelar placeboeffekten någon större roll för det kliniska förloppet hos dagens patienter i det västerländska samhället? Det finns exempel på väl utförda placebostudier, där man verkligen med rimlig vetenskaplig säkerhet har kunnat isolera en placeboeffekt. Resultaten av flera smärtstudier är ett exempel [2, 3].

När det gäller doktors roll i mötet med patienten finns åtminstone ett par välgjorda studier: dels från en vårdcentral i Sverige [4], dels från en engelsk allmänläkarmottagning [5]. Båda undersökningarna visade på en signifikant bättre symptomlindring dagarna efter konsultationen i de grupper som fått ett mer positivt bemötande än i grupperna som fått ett sämre bemötande. I hypertonisammanhang finns flera studier som visar blodtrycksförändringar av placebostimuli; en blodtryckssänkning demonstrerades bl a i en studie av Suchman och Ader [6].

Vad beträffar verbala stimuli (suggestion) har i en studie av två kvinnor med graviditetskräkningar kräkmedlet ipekuanas emetiska effekt kunnat vändas till sin motsats [7]. Vidare har inhalation av koksaltlösning hos astmapatienter både kunnat utlösa och häva ett astmaanfall beroende på den information patienten fått om den inhalede lösningens effekt [8]. I en annan studie kunde manifestationerna av kontaktallergi helt styras med suggestion [9].

I en undersökning av postoperativa besvär efter tandextraktion hos ett stort patientmaterial erhöles en signifikant, ca 30-procentig, minskning av både lokalt ödem och serum-CRP när patienterna trodde att de fått lokal ultraljudsbehandling, men då apparaturen i själva verket inte hade varit påslagen [10].

Problem och missuppfattningar

Problemen med placebobegreppet är flerfaldiga. Ett fundamentalt sådant är att det ännu så länge saknas en entydig, allomfattande definition av begreppet och vilka mekanismer som ger upphov till reaktionerna. Det råder visserligen en allmän acceptans av att symboliska stimuli och psykologiska mekanismer är verksamma när en placeboreaktion utlöses, men därefter upphör enigheten. Det finns ett spann mellan å ena sidan de som hävdar att man med ett placebostimulus endast menar en behandling som liknar en vanlig sådan [11] och å andra sidan de som anser att placebostimuli och placeboreaktioner förekommer i varierande grad hos patienter i alla sjukvårdssituationer, inte bara i förment terapeutiska sådana [12].

Evidensbaserad medicin

Ett annat problem är den stora floran av vetenskapligt undermålig litteratur på området, inte minst då det gäller att försöka isolera den rena placeboeffekten från andra interventioner och från tidseffekter, inkluderande regression mot genomsnittet. I många publikationer, bl a av randomiserade dubbelblindstudier med läkemedel, har placeboeffekten helt felaktigt likställts med den totala behandlingseffekten i placebogruppen.

Den amerikanske placeboforskaren Henry Beecher publicerade 1955 ett arbete som blivit mycket citerat, där han sammanställde resultaten i placebogrupperna från 15 slumpvis (?) utvalda läkemedelsstudier och fann att en förbättringseffekt förelåg i medeltal i ca 35 procent i dessa grupper [13]. Denna irrelevanta siffra är, förutom att den inte säger ett skvatt om den rena placeboeffekten, dessutom ett medelvärde från 15 studier och har sedan dess förlett många att tro att placeboeffekter skulle ha något slags inneboende konstans. I själva verket talar allt för att det är precis tvärtom, dvs att ett karakteristikum för placeboeffekter just är deras variabilitet, både mellan individer och för en enskild individ i olika situationer [14]. Denna stora, till synes nyckfulla, variabilitet har givetvis bidragit till svårigheterna att bedriva forskning kring placebo fenomenet.

En obesvarad fråga i placebosammanhang är också i vilken utsträckning placeboreaktioner är verksamma i den vardagliga sjukvården, dvs utanför den vetenskapliga provnings situationen. Det har av placeboforskare hävdats att placeboeffekterna är lägst i dubbelblindstudier, högre i enkelblindstudier och högst i okontrollerade kliniska rapporter [14]. Forskning saknas inom området, eftersom underlaget för slutsatser angående placeboreaktioner i den dagliga sjukvården grundar sig huvudsakligen på fallrapporter.

En metaanalys

Asbjörn Hróbjartsson och Peter Gøtzsche, verksamma vid avdelningen för medicinsk filosofi och klinisk teori vid Köpenhamns universitet och vid Nordiska Cochranecentret vid Rigshospitalet i Köpenhamn, publicerade en artikel i ett majnummer av New England Journal of Medicine med huvudrubriken »Is the placebo powerless?» [15]. Artikelnen har väckt stor uppmärksamhet såväl inom medicinska kretsar som i massmedierna och uppenbarligen lett till att man på vissa håll



har ifrågasatt existensen av placeboeffekter överhuvudtaget inom sjukvården [16].

Författarna utförde en metaanalys på 114 publicerade behandlingsstudier från 1946 till 1998 rörande diverse tillstånd, där man förutom en placebogrupp inkluderade en obehandlad kontrollgrupp. Resultaten vad beträffar positiva placeboeffekter vid denna analys var magra; inga statistiskt signifikanta skillnader mellan placebo- och kontrollgrupperna förelåg när det gäller binära parametrar. Däremot framkom positiva behandlingseffekter rörande kontinuerliga, subjektiva parametrar, statistiskt säkerställda för smärta.

Kan vi avföra placeboeffektens betydelse?

Kan vi därmed avföra placeboeffekternas betydelse i kliniska sammanhang? Knappast, åtminstone inte på grundval av denna metaanalys. Begränsningarna med den är för det första att man här valt ut endast publikationer där man använt sig av den snäva definitionen av ett placebostimulus, dvs placebo som en behandlingsintervention. Andra stimuli av symbolisk art, som t ex bemötandet i sjukvården, har därför inte beaktats.

För det andra är det inte möjligt att hävda att en obehandlad kontrollgrupp verkligen är obehandlad i strikt bemärkelse. Kontrollgruppen får en särställning bara genom att den blir uppmärksam och kontrollerad, och man har rimligen inte alltid full kontroll över eventuella andra interventioner som gruppen utsätts för under studiens gång. I allmänhet kan man inte tala om en sjukdoms eller ett tillstånd spontana förlopp i någon absolut bemärkelse.

Placeboreaktionernas uppenbara variabilitet torde också medföra att man inte kan erhålla något konsistent terapieresultat vid jämförelser mellan en placebo- och kontrollgrupp i olika studier.

Ytterligare kritiska kommentarer rörande Hróbjartssons och Götzsches arbete och slutsatser framkom i en ledarkommentar i samma tidskriftsnummer [17] och senare i insändare i samma tidskrift [nr 17 den 25 oktober 2001]. Man berörde bl a begränsningarna med metaanalyser i största allmänhet, risken för typ 2-fel, den dåliga metodologiska kvaliteten i en del av de inkluderade studierna, oklarheter vad gäller både upplysningar och instruktioner till patienterna i de olika grupperna.

De inledningsvis citerade exemplen gör det svårt att förneka att det förekommer placeboeffekter, positiva som negativa, i sjukvården, även om det är svårare att mäta deras storlek i den kliniska vardagen. Inom vissa kretsar har det måhända ställts överdrivna förväntningar på positiva placeboeffekter.

Kanske är det så som den amerikanske gastroenterologen Howard Spiro har uttryckt det: placebo kan minska lidande, men botar inte sjukdomar [18].

Man borde kanske också mera försöka synliggöra eventuella noceboeffekter i sjukvården. Om man vidgar perspektivet något finns exempel på symboliska händelser i livet också utanför sjukvårdens domäner som kan påverka människors hälsotillstånd. Mycket talar t ex för att stress i form av negativa livshändelser (t ex skilsmässa, arbetslöshet) medverkar till utlösande av Graves' tyreotoxikos [19]. Samma dag som jordbävningkatastrofen i Los Angeles inträffade år 1994 noterades inom regionen 24 fall av plötslig död beroende på aterosklerotisk kardiovaskulär sjukdom, mot i genomsnitt 4,6 fall per dag veckan före [20].

Frågan i rubriken till denna översikt anser vi bör besvaras med ett ja, men kunskaperna kring placebofenomenet är fortfarande bristfälliga. Det finns därför behov av mera forskning på området, i förhoppningen om att kunna förbättra omhändertagandet av sjuka människor.

Referenser

- Shapiro AK. A contribution to a history of the placebo effect. *Behav Sci* 1960;5:109-35.
- Boureau F, Leizorovicz A, Caulin F. Effet placebo sur les douleurs métastatiques osseuses. *Presse Med* 1988;17:1063-6.
- Laska E, Sunshine A. Anticipation of analgesia. A placebo effect. *Headache* 1973;13:1-11.
- Olsson B, Olsson B, Tibblin G. Förväntanseffekter vid tonsillitbehandling. Personligt bemötande gav snabbare tillfrisknande. *Läkartidningen* 1988;85:610-3.
- Thomas KB. General practice consultations: is there any point in being positive? *BMJ* 1987;294:1200-2.
- Suchman AL, Ader R. Classic conditioning and placebo effects in crossover studies. *Clin Pharmacol Ther* 1992;52:372-7.
- Wolf S. Effects of suggestion and conditioning on the action of chemical agents in human subjects. *The pharmacology of placebos. J Clin Invest* 1950;29:100-9.
- Luparello T, Lyons HA, Bleeker ER, McFadden ER. Influences of suggestion on airway reactivity in asthmatic subjects. *Psychosom Med* 1968;30:819-25.
- Ikemi Y, Nakagawa S. A psychosomatic study of contagious dermatitis. *Kyushu Journal of Medical Science* 1962;13:335-50.
- Hashish I, Hai HK, Harvey W, Feinmann Ch, Harris M. Reduction of postoperative pain and swelling by ultrasound treatment: a placebo effect. *Pain* 1988;33:303-11.
- Kienle GS, Kiene H. Placebo effect and placebo concept: a critical, methodological and conceptual analysis of reports on the magnitude of the placebo effect. *Altern Ther Health Med* 1996;2:39-54.
- Peck C, Coleman G. Implications of placebo theory for clinical research and practice in pain management. *Theor Med* 1991;12:247-70.
- Beecher HK. The powerful placebo. *JAMA* 1955;159:1602-6.
- Shapiro AK, Shapiro E. The placebo: is it much ado about nothing? In: Harrington A, editor. *The placebo effect. An interdisciplinary exploration*. Cambridge, Massachusetts: Harvard University Press; 1997. p 12-36.
- Hróbjartsson A, Götzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* 2001;344:1594-602.
- McDonald CJ. To the editor. *N Engl J Med* 2001;345:1276-7.
- Bailard JC. The powerful placebo and the wizard of Oz. *N Engl J Med* 2001;344:1630-2.
- Spiro H. Clinical reflections on the placebo phenomenon. In: Harrington A, editor. *The placebo effect. An interdisciplinary exploration*. Cambridge, Massachusetts: Harvard University Press; 1997. p 37-55.
- Winsa B, Adami HO, Bergström R, Gamstedt A, Dahlberg PA, Adamson U, et al. Stressful life events and Graves' disease. *Lancet* 1991;338:1475-9.
- Leor J, Poole WK, Kloner RA. Sudden cardiac death triggered by an earthquake. *N Engl J Med* 1996;334:413-9.

Kommunikativa problem i den evidensbaserade medicinen

II Under senare hälften av 1900-talet har ett samhällsvetenskapligt intresse för medicinsk vetenskap och medicinsk teknik vuxit sig allt starkare. Forskningen har mer och mer inriktats mot studiet av förmedling av medicinsk specialistkunskap. Antropologerna som befinner sig inom detta forskningsfält genomför ofta sina studier med ett intresse för interaktion och kommunikation. Forskning kring budskapsförändringar mellan avancerade medicinska forskningsresultat och klinisk information har blivit ett prioriterat område inom medicinsk antropologi. Sådan forskning ger nya perspektiv på hur medicinsk kunskap översätts från specialiserade forskningsresultat till såväl det kliniska mötet som vidare ut till en icke-specialiserad allmänhet via media. Man har också intresserat sig för hur sådan kunskap tolkas och används i en enskild människas liv. I detta sammanhang förefaller diskussionen kring evidensbaserad medicin vara självklart intressant för en antropolog.

Enligt tidigare artiklar i Läkartidningens serie är grundtanken med evidensbaserad medicin att »tillgängliga vetenskapliga resultat skall komma till direkt användning i det praktiska kliniska arbetet« [1, 2]. Ytterligare ett syfte med evidensbaserad medicin är att klinisk medicin skall bygga på bästa möjliga vetenskapliga underlag integrerat med den enskilde läkarens kliniska kunskap och erfarenhet [3] eller att det skall vara »en praktisk filosofi, som innebär att varje åtgärd i sjukvården skall vara baserad på bästa möjliga tillgängliga kunskap« [4]. Många är rösterna kring detta i Läkartidningens serie, till vilken skara jag nu sällar mig med ett inlägg där fokus ligger på en annan nivå – den tolkande nivån.

Sett från min utgångspunkt som samhällsvetare/humanist finns en komplexitet i evidensbaserad medicin som påverkar kommunikationen kring de vetenskapliga resultaten. Det som burit mitt forskningsintresse framåt genom åren är utvecklingen av den medicinska teknologin och dess resultat inom den form av hälso- och sjukvård som det moderna samhället kräver. Forskningen har fokuserat två extremer, nämligen medicinskt hanterande av å ena sidan människor med mycket stort lidande utan patologisk förklaring och å andra sidan patologiska fynd eller riskdiagnoser hos helt »friska« människor. De studier jag har gjort kring läkemedelsanvändning, dess symbolvärden och placebo-/noceboeffekter [5-7]; prevention, hälsokontrollerande verksamhet, kolesterol-screening [8-11]; kroniskt trötthetssyndrom, lidande utan pa-

SAMMANFATTAT

Det finns en komplexitet i evidensbaserad medicin som skapar problem i kommunikationen kring de vetenskapliga resultaten.

Den kommunikationsteoretiska utgångspunkten för interaktion mellan människor är att parterna tillsammans skapar innehållet i samtalet. Det är därför fel att tala om en sändare och en mottagare av ett budskap när människor möts i vårdssituationer.

Information är aldrig neutral och kan därför inte lämnas över från en person till en annan i ett opåverkat skick. Den tolkas alltid och blir därför omöjlig att kontrollera för den som lämnar informationen.

Ansvaret vilar tungt på den som har som sin uppgift att förmedla budskap om viktiga och ibland livsavgörande frågor oavsett om de bygger på säker eller osäker evidens.

Det krävs tidig träning i ett kommunikationstänkande där man som läkare blir varse samtalets centrala plats som verktyg i all hälso- och sjukvårdsverksamhet.

Evidensbaserad medicin

tologi [12-14]; och till sist genetisk kartläggning, riskdiagnoser vid familjär cancersjukdom [15-18], speglar just dessa två ytterligheter som skall hanteras i vården. Det kliniska mötet, kommunikationen, informationen och hur den förmedlas, tolkas och förstås av de inblandade parterna, fokuseras i dessa studier. Samtliga analyser visar entydigt att patienterna har svårt att tolka vetenskapliga grunder för diagnos, bedömning och behandling. Analyserna påvisar också de svårigheter som ligger i själva informationen när vetenskaplig kunskap skall översättas till en klinisk situation för att en enskild person

Information är aldrig neutral. Den tolkas alltid på människors egna villkor utan att kunna kontrolleras av den som lämnar informationen. Detta är en del i komplexiteten i medicinsk kommunikation. Bilden är arrangerad, och personerna har inget samband med artikeln.

skall kunna ta beslut och leva vidare med en ny insikt om sitt tillstånd och sina möjligheter.

Men i de speciella medicinska sammanhang som jag har studerat är det inte endast information och interaktion som innebär problem, osäkerhet kan också uppstå i de fall där den kliniska verksamheten måste drivas samtidigt som forskningen pågår. I den situationen finns ännu inte någon beprövad erfarenhet. Den medicinska teknologin ligger steget före, men kommunikation kring »resultat« sker ändå. Även om dessa osäkra resultat med nödvändighet måste översättas till direkt klinisk verksamhet och tolkas av patienten utan att osäkerheten upptäcks, finns risker för patienten vilka kan komma att uppenbaras i efterhand. Detta gäller exempelvis informationen om bieffekter av vissa läkemedel, informationen om när förhöjda kolesterolvärden får ogynnsamma effekter, informationen om kroniskt trötthetssyndrom och diagnoser där avsaknaden av fynd inger såväl lättnad som besvikelse hos patienten – diagnosen ger inget hopp om bot. Det gäller slutligen information om de genetiska anlagens betydelse för cancersjukdom, där forskningen hela tiden leder till nya upptäckter.

Hela familjer, hela släkter kan medikaliseras i och med informationsamtal med en enskild person. Enligt mina studier hanteras dessa medicinska frågor i klinisk verksamhet med utgångspunkt i osäker evidens.

Parterna skapar samtalet tillsammans

Alla som arbetar med allvarliga sjukdomstillstånd och riskinformation i samband med medicinska beslut vet att det inte bara gäller att hitta sätt att beräkna och bedöma risker utan också att hitta sätt att presentera dem på för dem de angår. Mätvärden skall anges, genetiska eller molekylära samband tydliggöras och bilder av kroppens inre tolkas och förklaras. Det huvudsakliga informationsarbetet utgörs av samtal. Medicinere och andra vetenskapsmän kan numera utföra sofistikerade test och göra beräkningar och mätningar för att noggrant förutse risker på ett generellt plan, men i mötet och samtalet – t ex i samband med en livshotande eller allvarlig sjuk-

dom eller skada – krävs kommunikativ kompetens av båda deltagarna [19].

Den kommunikationsteoretiska utgångspunkten för interaktion mellan människor är att parterna tillsammans skapar innehållet i sitt samtal. Det är därför fel att tala om en sändare och en mottagare av ett budskap när människor möts i vårdsituationer. Detta innebär att information aldrig är neutral och kan lämnas över från en person till en annan i ett opåverkat skick. Den tolkas alltid och blir därför okontrollerbar för den som lämnar informationen. Ansvar vilar tungt på den som har till sin uppgift att förmedla budskap om viktiga och ibland livsavgörande frågor, oavsett om de bygger på säker eller osäker evidens.

Medicinska praktiker är inte alltid medvetna om att deras information omtolkas, och kanske förvanskas, till det som är möjligt att förstå för patienten. Dessutom är det faktum att information alltid tolkas på människors egna villkor bara en del i komplexiteten.

Patientens rättigheter ställer krav på informationen

Dessa teoretiska resonemang väcker tankar om de problem som till exempel den nya lagstiftningen kring människors inflytande över och deltagande i val av behandling och möjlighet till förnyad medicinsk bedömning kan komma att föra med sig. Här handlar det om att se över huruvida lagstiftarens intentioner fått genomslag i praxis. En sådan fråga är med utgångspunkt i det ovan skisserade utomordentligt komplicerad.

Kraven som införts för att stärka patientens ställning har sedan den 1 januari 1999 ställts på vårdgivaren i Hälso- och sjukvårdslagen [20] och på hälso- och sjukvårdspersonalen i Lagen om yrkesverksamhet på hälso- och sjukvårdens område [21]. Det gäller frågor som:

- individuellt anpassad information,
- möjlighet att välja behandlingsalternativ,
- möjlighet att få förnyad medicinsk bedömning.

Vad vet vi om när patienter t ex vågar begära en alternativ medicinsk behandling? Vissa omständigheter som har med själva vårdkulturen att göra har också avgörande betydelse vid

möten mellan människor där allvarliga beslut ska tas och där den lagstiftade rätten som patienten har att få en *second opinion* skall introduceras:

- Yrkes stolthet och vana vid expertrollen från läkarens sida är sådant som kan göra att frågan om *second opinion* uteblir.
- En upplevd väl fungerande kommunikation mellan patient och läkare inom vårdkulturens etablerade ramar kan också göra att *second opinion* känns onödig för båda parter.
- En patients oro över att själv behöva välja vad som är bra eller dålig behandling kan också göra att möjligheten till *second opinion* inte kommer i fråga – vårdkulturen har också här givit utgångspunkter för hur man ska agera och tänka. Beroendeställningen fungerar också som ett skydd och en trygghet för patienter.

Neutral information – en myt

I olika länder i västvärlden gör man alltmer idoga försök att skapa konsensus kring hur patienten skall informeras i samband med diagnos och behandling vid svåra sjukdomstillstånd. The American Board of Medical Genetics kräver av sina medlemmar inte bara att de skall bli skickliga i att intervjua så att de »från patienten eller familjen kan bringa i dagen (elicit) den information som är nödvändig för att komma fram till den riktiga (appropriate) slutsatsen« [22], de måste också »överföra (transmit) relevant (pertinent) information effektivt (dvs på ett sätt som är meningsfullt för individen eller familjen)«. I dessa citat kan vi se att läsaren inbjuds att se information som något som kan »bringas i dagen« och »överföras«. Det är mycket vanligt att vi tänker på hur det går till att informera med hjälp av en överföringsmetafor. Information ses då som något väl definierat och avgränsat som kan formuleras av avsändaren och orört överförs till mottagaren. Formuleringarna från The American Board of Medical Genetics antyder ett ideal av neutral överföring av fakta. Men samtidigt pekar användning av adjektiven »relevant« och »riktig« på insikten att information alltid tolkas och värderas från ett visst perspektiv och i ett visst sammanhang.

Den neutrala informationen är således en myt [23]. Rådgivning är inte enbart en överföring av information från en som vet till en som inte vet. Den är i själva verket en tvåvägsprocess genom vilken båda deltagarna lär sig mer om sig själva och varandra och om det man talar om. De som får råd tar inte till sig nakna fakta. De tolkar fakta i enlighet med sin tidigare förståelse, sina värderingar och sina känslor och lär sig det som berör just dem. Fakta är fakta-för-dem, fakta som laddas med deras egna livserfarenheter och känslor. De kan uppfattas och tolkas så att de förorsakar stark ångslan och känslor av hjälplöshet, men också så att de medför lättnad och stor handlingskraft [10, 17]. I läkare-patient-samtal talar dock inte sällan läkaren om information som om det rörde sig om neutrala fakta och om informationsutbyte som om vore det »bara« fråga om en överföring av fakta från en person till en annan. Begreppet »informed consent« är bara ett av de besvärliga kommunikationsfrågor som här leder till problem.

I rådgivnings- och informationssamtal som gäller å ena sidan förhöjda kolesterolvärden och å andra sidan ärftlig cancer talade man om »risk« som något en person kan »ha«. Men för att närmare förklara risken används siffror av olika slag: testvärden, intervall, gränsvärden eller procenttal. I många läkarsamtal spelar alltså siffror en avgörande roll. Men frågan är hur lätt det är att förstå dem. Hur kommunicerar man kring epidemiologiska och genetiska riskbedömningar – översätta till gränsvärden eller matematiskt uttryckta sannolikheter – i det kliniska mötet mellan läkare och lekmän? Hur uppfattas

och omtolkas sådana riskberäkningar i det kliniska mötet och i individens livsvärld?

Professionell osäkerhet i kommunikation kring testresultat

Avsikten med det preventiva arbetet inom hälsovården är att spåra och undanröja riskfaktorer för ohälsa i befolkningen. För att kunna göra sådana upptäckter tar hälso- och sjukvården initiativ och gör undersökningar bland en befolkning som själv betraktar sig som frisk. Test av detta slag har en aura av precision och vetenskaplig objektivitet som förstärker deras trovärdighet. De utgör ofta ramen för ett professionellt och populärt förhållningssätt kring sociala och individuella problem som i sin tur påverkar verksamheter inom skilda institutioner och social policy. Som exempel kan nämnas testning av kolesterolvärden. En individs värden kan vara föremål för diskussion inom den medicinska expertisen, men också hemma vid matbordet och tillsammans med vänner.

Kolesterolets betydelse som riskfaktor för sjuklighet och dödlighet i hjärt-kärlsjukdom är en kontroversiell fråga som skapar osäkerhet inom den medicinska professionen. Denna så kallade professionella osäkerhet framträder vid tillstånd där vårdbehoven varken är akuta eller absoluta och där det inom läkarkåren råder olika uppfattningar om nyttan och riskerna med de interventionsmöjligheter som står till buds. Kontroversen omfattar i detta fall ett antal frågor, bl a huruvida förhöjda kolesterolvärden utgör en riskfaktor för sjukdom och inte en sjukdom i sig, om det är etiskt acceptabelt att genom masshälsoundersökning av hela befolkningsgrupper ställa en (sjukdoms)diagnos på symptomfria individer, vid vilken kolesterolnivå behandling är indicerad samt nyttan av och riskerna med olika typer av behandling.

I uppföljningssamtal bland 12 av 28 40-åriga män som ingick i en studie om hälsoinformation och som fått information om att deras skadliga kolesterol var förhöjt visade det sig att själva värdet, exempelvis 6,8, var fokus för deras uppmärksamhet [10]. Männerna hade inte bara idéer om vikten av att veta vad medelvärdet var, de hade också föreställningar om och ville få reda på det ideala testresultatet. Några föreställde sig att det skulle vara 0, andra hade andra förslag. När sjuksköterskan upplyste en av männen om att han på kolesteroltestet hade 6,7 »och det ligger ju precis här i övre gränsen, det är bra om man ligger under 6,8 och det gör ju du, men du ligger ju ganska högt upp ändå, nära gränsen« så ville han veta mer exakt var man skall ligga. Han föreslog frågande: »ska man ligga på 5 eller?«

Som exemplet visar uttryckte mannen en stark önskan att få veta vad som är det »ideala«, »genomsnittliga« eller »förväntade« resultatet.

Männen ville överlag vara normala. Förutom en önskan att stadfästa den egna normaliteten så visar männens frågor hur svårt de har att tolka siffror utan sammanhang. Även om ett fåtal, precisa mätvärden – t ex längd och vikt – nästan framstår som en del av vår identitet, är andra tillstånd eller kroppsliga egenheter oftast inte begripliga i matematiska termer. Vi har några ganska vaga, generella begrepp och tenderar att uppfatta höga tal i samband med kroppen som mer riskabla än låga. I termer av individuell hälsa är ett högt blodtryck eller en hög kolesterolnivå negativ. Men med några mätvärden, t ex vissa blodberäkningar, är det tvärtom. Här är det bättre att ha ett högt värde. Det som komplicerar uppfattningen och tolkningen av det personliga mätvärdet är det faktum att värdet egentligen talar om en epidemiologiskt framtagen siffra som handlar om ett medelvärde hos en population. Det blir aldrig möjligt att vara helt säker vid avgörandet huruvida en viss siffra är bra eller dålig hos en enskild individ. En av männen framförde förslaget att »jag kanske är en person som har ett eget värde, en gräns som inte är farlig«. På sätt och vis kan man säga att denne man valde att inte vara med i det statistis-

ka sammanhanget. Han tolkade helt enkelt sina kolesterolvärden inom ramen för sin egen livsvärld. Vi har nämligen inte så lätt – som de unika individer vi ju faktiskt är – att tänka på oss själva som statistiska typexempel.

I efterintervjuerna berättade sjuksköterskan att hon var väl medveten om den professionella osäkerhet som råder kring värdena. Forskarna kan t ex inte påstå att det är bättre att ha ett lågt testvärde än ett högt eller att det i någon mening är bättre att ha 4,2 på kolesteroltestet än 5,8. Det innebär att en kliniskt applicerad riskmodell inte gör hierarkiskt grundade skillnader mellan de testresultat som faller under gränsvärdet.

Svårt att informera patient inför eget beslut

Ett exempel från ett möte mellan läkare och frisk kvinna när det gäller ärftligt anlag för äggstockscancer visar hur svårt det är med information när kraven ställs på att patienten själv skall ta beslutet [17, 24]. Läkaren går igenom de alternativ som finns för att kvinnan skall kunna avgöra hur hon vill göra. Dessa är, enligt honom, att testas regelbundet eller att genomgå en profylaktisk operation. Det innebär att man tar bort hennes äggstockar. Läkaren bemödar sig om att göra klart för kvinnan att hon måste fatta sitt beslut själv och säger att han »bara« kan ge henne den information som han har tillgång till. Ordet »bara« kan antyda en syn på information som enbart fokuserar den kognitiva aspekten – ett rent tankemässigt innehåll som i en väderleksrapport. Att informera blir då detsamma som att meddela några odiskutabla fakta. Även om läkaren lyckas med att presentera sig själv som neutral och som om kvinnan har ett fullständigt fritt val av handling, kan hon mycket väl tolka den neutrala formulerade informationen som ett metabudskap med innebörden: »Du borde genomgå operationen.« (Efterintervjuerna visar att så var fallet med just den här kvinnan.) [24] Det är också känt att när någon informerar oss om en tänkbar lösning på ett problem tolkar vi ofta detta som ett råd eller en instruktion – dvs som en handlingsplan.

Informationen om risk för ärftlig sjukdom via genealogiska kartor ersätts av eller kompletteras alltmer med information via DNA-utsnitt i och med att man finner mutationer i gener som går att bekräfta via blodprov. Denna form av information baserar sig på evidens som är konkret och påtaglig och kan ses av patienten tillsammans med läkaren.

Genom den snabbt utvecklade medicinska teknologin skapar forskare ett bildmaterial av kroppens mikrovärld. Man kan anta att visualiseringen av skeenden som teknologin ger möjlighet att belägga förser forskarna och deras medhjälpare med ett speciellt språk, med vilket de kommunicerar med varandra och med lekmän som berörs. Hur används språket för att representera kroppens inre bilder tillsammans med patienten? Hur tolkar och förstår personer med svåra sjukdomar, som måste behandlas med riskfyllda behandlingar, siffrorna eller bilderna av kroppens status? Hur förstår man risk för genetiskt betingad sjukdom via visualiserande av atypisk DNA eller liknande bilder av kroppens inre förhållanden? Krävs här en speciell förmåga att tolka abstrakta bilder?

Att siffror, bilder och sannolikhetsberäkningar används är givetvis ett sätt att försöka ge korrekt information från den medicinska expertisens sida. Man kan också säga att den retoriska effekt som eftersträvas är just »vetenskaplighet«, dvs evidensbaserad kunskap. Beskedet skall ha tyngd och sanningsvärde, och siffror är nog den tyngsta delen i den retoriska arsenal som läkarna har till sitt förfogande. Men siffervärden och bilder av kroppens inre tolkas av dem som hör och ser dem vid varje enskilt tillfälle och på deras egna villkor.

Träning i kommunikation bör ges i läkarutbildningen

Problemen som jag beskrivit i denna artikel är mest fokuserade på kommunikation mellan människorna som tillsammans

deltar i vårdsituationer. Men användandet av evidensbaserad medicin i vårdsituationer innebär inte endast svårigheter att förmedla information. Det finns också i dessa sammanhang medicinsk verksamhet, exempelvis när man ger »riskdiagnoser« vid ärftlig cancer, där information om just risker blir oklar och felaktig på grund av osäkerhet vad gäller översättning av vetenskaplig kunskap till den kliniska situationen [1, 2, 25]. Detta kan ofrivilligt komma att leda till mer skada än nytta. När vi som patienter informeras om olika risker vad gäller såväl sjukdomsprocesser som behandlingsresultat vid allvarlig sjukdom har vi, oavsett hur vi tolkat informationen, sällan en direkt benägenhet att ifrågasätta.

I min forskning samt i mina samtal med läkare som besitter den nya och mycket avancerade kunskap som exempelvis vuxit fram genom molekylärbioforskning finns beskrivningar av den frustration som relationen forskning–praktik ofta leder till. Varje dag medför nya framsteg inom forskningen, men klinisk erfarenhet släpar efter. Tanken på hur denna mycket specialiserade kunskap skall kunna översättas till patienten är pressande. Samtidigt som man inser hur mycket det kan betyda att läkemedel i framtiden kan skraddarsys och olika sjukdomar stoppas som resultat av pågående forskning, förefaller relationen mellan kunskapen och det kliniska mötet mer och mer svårhanterlig.

I en artikelserie i LT kring framtidens läkare finns en stark förändringsvilja och en, enligt mitt sätt att se, nydanande fräschör i tonen kring den egna situationen [26]. Men många goda viljor krävs för det praktiska genomförandet. En möjlig väg till en framtida läkargärning med mindre stress kunde till exempel vara tidig träning i ett kommunikationstänkande under själva läkarutbildningen. Under en god sådan träning parallellt med medicinska studier kunde man som blivande läkare bli varse och ta på allvar samtalets centrala betydelse i all hälso- och sjukvårdsverksamhet. Medvetenheten om att samtal är något parterna skapar tillsammans och som alltid sker på olika villkor kan utgöra en grund för större lyhördhet och anpassning i interaktionen mellan läkare och patient. En förstärkt kunskap om att just samtalet är det verktyg som hjälper patienten bäst skulle på sikt kunna utveckla en förtrogenhet med att informera om snåriga bevisbördor kring avancerade forskningsresultat på patientens villkor. Häri ligger en av de största frågorna med den evidensbaserade medicinen.

Den kommunikativa problematik som ryms i dagens hälso- och sjukvård kräver mycket starka viljor och en oändlig mängd fantasi för att smälta samman en ambition till »patientens rätt« och evidensbaserad medicin på erfarenhetens grund.

Referenser

1. Taube A, Malmquist J. Räkna med vad du tror. Bayes' sats i diagnostiken. Läkartidningen 2001;98:2910-3.
2. Taube A, Malmquist J. Räkna med vad du tror. Bayes – inte P-värdet – mäter tilltron. Läkartidningen 2001;98:3208-11.
3. Eliasson M. Evidensbaserad medicin »begins at home«: incitament för bättre beslutsunderlag finns i den kliniska vardagen. Läkartidningen 2000;97:3860-5.
4. Werkö L. Ett konkret redskap för vårdens verkstadsgolv. Läkartidningen 2001;22:2710-3.
5. Sachs L. Misunderstanding as therapy. Culture, medicine and psychiatry 1989;13:335-49.
6. Sachs L. Is there a pathology of prevention? Culture, medicine and psychiatry 1995;19:503-25.
7. Adelswärd V, Sachs L. Den preventiva medicinens dilemman. Läkartidningen 1994;91:3524-9.
8. Adelswärd V, Sachs L. The meaning of 6.8: numeracy and normality in health information. Soc Sci Med 1996;43:1179-87.
9. Sachs L, Evengård B. Speciell patient-läkarerelation vid kroniskt trötthetssyndrom. Läkartidningen 2000;97:176-81.

14. Hydén LC, Sachs L. Suffering, hope and diagnosis: on the negotiation of chronic fatigue syndrome. *Health* 1998;2:175-93.
15. Sachs L. Att leva med risk: fem kvinnor, gentester och kunskapens frukter. Stockholm: Gedins förlag; 1998.
16. Adelswärd V, Sachs L. Risk discourse: recontextualization of numerical values. *Text* 1998;18:191-210.
17. Sachs L. Knowledge of no return: getting and giving information about genetic risk. *Acta Oncol* 1999;38:735-41.
19. Sachs L. Att tolka budskap i vårdmöten över kunskapsgränser. *Läkartidningen* 2000;97:3752-8.
20. Hälso- och sjukvårdslagen. SFS 1982:763.
22. Brock SC. Narrative and medical genetics: on ethics and therapeutics. *Qual Health Res* 1995;5:150-68.
23. Markova I. Medical Ethics: a branch of societal psychology. In: Himmelweit H, Gaskell G, editors. *Societal Psychology*. London: Sage, 1990.
24. Adelswärd V, Sachs L. Budbärarens dilemma: om moraliska aspekter på information. *Socialmedicinsk tidskrift* 2001;5:406-16.
25. Sachs L, Taube A, Tishelman C. Risk in numbers: difficulties in the transformation of genetic knowledge from research to people: the case of hereditary cancer. *Acta Oncol* 2001;40:445-53.
26. Adami J, Lagergren J, Tsai L. Hur ska vi återta professionalismen i läkaryrket? *Läkartidningen* 2001;98:4480-3.

I Läkartidningens elektroniska arkiv
<http://larkiv.lakartidningen.se>
är artikeln kompletterad med fullständig referenslista.

SUMMARY

Communication problems in evidence-based medicine

Lisbeth Sachs

Läkartidningen 2002;99:757-62

From a humanistic, social scientific perspective, the most complex task in evidence-based medicine lies in the communication of specialized medical knowledge to non-professionals.

Information is never simply the neutral transmission of facts, not even when dealing with scientific knowledge and research. It is always interpreted and evaluated from a particular perspective in a specific context.

That information can be neutral is thus a myth. In all medical consultations the process of communication is not just a matter of transmitting information from one who knows to one who does not. Knowledge created and formulated in a scientific context is thus recontextualised first in a clinical situation and then as an interpreted version in people's real lives.

Furthermore there are difficulties when practice must be based on current research, in a situation in which no prior clinical experience exists and in which results are interpreted and used regardless of the relative certainty of current evidence.

Correspondence: Lisbeth Sachs, Tema Kommunikation, Linköpings Universitet, SE-581 83 Linköping, Sweden. (lisbeth.sachs@swipnet.se)

Hur man skriver en vetenskaplig artikel



– och får den publicerad!

Läkartidningen och **BMJ** (British Medical Journal) arrangerar tre endagskurser för medicinska skribenter: 22, 23 och 24 april 2002.

Kursen presenterar hur man på ett effektivt sätt omsätter tankar och koncept till en välskriven uppsats – hur man organiserar och analyserar sitt material, formulerar slutsatser och hur man korresponderar med redaktioner och redaktörer enligt »spelets regler«. Rent praktiskt bygger kursen på att man med utgångspunkt från sitt eget material skriver en artikel under kursdagen. Kurserna vänder sig därför i första hand till forskarstuderande med färdiginsamlade och bearbetade data.

Kursledare är Tim Albert, vetenskapsjournalist på BMJ samt Harvey Markovitch, chefredaktör på **Archives of Disease in Childhood**. Kurserna har gått flera år i Storbritannien och rönt stor uppskattning.

Kursavgiften beräknas vara 2 700 kr.

Så här anmäler Du Dig. Ingen särskild anmälningsblankett behövs. Skicka in fullständiga adressuppgifter, registreringsdatum till forskarutbildningen, ett CV på högst 300 ord samt en kort motivering varför Du vill gå kursen till: inga-maj.lagerholm@lakartidningen.se

Din anmälan måste vara hos oss senast den 1 mars.

Läkartidningen vill genom kursen stimulera intresset för att skriva vetenskap på svenska, och deltagare intresserade av att medverka som skribenter av vetenskapliga referat har företräde.

Välkommen med Din anmälan
– antalet platser är begränsat!

Josef Milerad
docent, medicinsk chefredaktör

Per Dalén, med dr, docent, Viken (pdalen@algonet.se)

Är placebotänkandet evidensbaserat?

Frågan om placeboeffektens vetenskapliga status har blivit aktuell på senare år framför allt genom två publikationer. Den första var en liten bok av Gunver Sophia Kienle [1], där Henry K Beechers tongivande artikel »The powerful placebo« från 1955 [2] granskades med ganska uppseendeväckande resultat. Beecher hade startat något av en vandringsägen genom sin slutsats att placebo genomsnittligt verkar i 35 procent av fallen, oavsett diagnos. Han lade också i praktiken fast uppfattningen att placeboeffekten är lika med andelen förbättrade i placebogruppen i en klinisk prövning. Den kritik som av och till framfördes på dessa punkter fick inget påtagligt genomslag under de 40 år fram till 1995, då Kienles bok kom ut [3-4]. Beechers artikel citerades ständigt; hans syn på saken hade gått in i det allmänna medvetandet.

Granskningar som avslöjade många brister

Kienles systematiska granskning avslöjade många brister och feltolkningar. Bl a bortsåg Beecher helt från de behandlade tillståndens naturförlopp och vissa andra faktorer som bidrar till att den förbättring som registreras i placebogruppen blir systematiskt grovt missvisande som mått på en »verklig« placeboeffekt.

Kienle rörde vid något ganska ömtåligt, och hon har givetvis i sin tur fått kritiker, som ofta tycks ha svårt att erkänna betydelsen av hennes insats. Alla som följt Beecher i fotspåren och okritiskt fört hans misstag vidare träffas emellertid av Kienles kritik.

Den andra stora händelsen på området inträffade i maj 2001, då Asbjørn Hróbjartsson och Peter C Gøtzsche i Köpenhamn publicerade sin metaanalys av 114 studier där både en placebogrupp och en obehandlad kontrollgrupp ingick [5]. Jämförelser med obehandlade kontroller är ju den enda direkta vägen att systematiskt studera den »verkliga« placeboeffekten. Deras resultat visade inga signifikanta placeboeffekter i studier som använde objektiva eller binära mått på utfall, däremot små signifikant positiva effekter vid smärtbehandling eller när effekten skattades subjektivt i en kontinuerlig skala.

Den kände cancerstatistikern John C Bailar skrev sin ledarkommentar i samma nummer av *New England Journal of Medicine* under rubriken »The powerful placebo and the wizard of Oz« [6]. Bailar inleder så här:

»Some myths really ought to be true. We react with surprise and pleasure when we encounter them and then believe them when they neatly and comfortably help to explain some confusing aspect of our world. Thereafter, evidence against them is unwelcome and not to be trusted. But some such myths are flawed and misleading.«

Ledaren utmynnar i kommentaren att placebo inte helt kan

SAMMANFATTAT

Placeboeffekten existerar, men det saknas evidens för att den är så universell och kraftfull som påstås i traditionen.

Debatten har nu för första gången förutsättningar att leda fram till en vetenskapligt balanserad syn på ett fenomen som länge betraktats som mer eller mindre magiskt.

Evidensbaserad medicin

avföras från diskussionen. Materialet är inte idealiskt, men en viss mindre placeboeffekt på ett så betydelsefullt område som smärta kunde dock påvisas.

Oavsett vad man kan säga om artikeln av Hróbjartsson och Gøtzsche har de i stort sett använt det material som finns att tillgå. Det betyder att ingen inom överskådlig tid kan lägga fram nya motsvarande data och dra helt andra slutsatser.

I Sverige har SBU utkommit med två böcker som direkt behandlar placebofrågan. Den första, »Patient-läkarrelationen: Läkekunst på vetenskaplig grund« [7], skrevs kollektivt av en författargrupp under ordförandeskap av Jan-Otto Ottosson. Här behandlas Beechers artikel utan särskilda reservationer på sid 79, och placeboeffekter på upp till 89 procent citeras ur litteraturen. Tio sidor senare konstateras att samma artikel »har utsatts för nedgörande kritik« av Kienle och Kiene [4], men läsaren får närmast intrycket att den kritiken utgjort ett villospår. I själva verket utgår Kienle och bokens författare från samma principer vid analysen av den skenbara placeboeffekten. Det hade varit följdriktigt om man i boken tydligt hade distanserat sig från den mycket omfattande litteratur som helt accepterar Beechers synsätt.

Författargruppen konstaterar att de »jämförande placebo-studier där en obehandlad kontrollgrupp ingått visar dels att det finns en sann placeboeffekt, dels att placeboeffektens storlek beror av patientens och läkarens gemensamma förväntningar samt olika aspekter av patient-läkarrelationens kvalitet«. Underlaget för dessa slutsatser redovisas inte entydigt, men sex studier med smärtbehandling ingår sannolikt.

I bokens sammanfattning påstås att »förbättringen [efter placebo] kan ibland kvarstå under lång tid«. För detta viktiga

påstående finns inga andra belägg än ett omnämmande av en artikel från 1958, som skall ha visat att effekter på smärta »kan vara bestående när behandlingen upphör«. Betydelsen av denna uppgift är naturligtvis omöjlig att värdera.

Nästa skrift från SBU heter helt enkelt »Placebo« [8]. Nio författare står här själva för varsitt kapitel. Harry Boström talar om Beechers artikel som »en klassiker« som »bekräftats i mycket omfattande studier«. Erik Hägg är fullt medveten om de problem som Kienle fört upp på dagordningen. Han drar överraskande pessimistiska slutsatser när det gäller att separera den »sanna« placeboeffekten från de olika andra faktorer som ingår i den skenbara effekten. Han är också tveksam till värdet av att jämföra placebobehandlade med obehandlade kontrollgrupper. Svårigheterna att bedriva invändningsfri placeboforskning måste i så fall bli stora.

Kienle kritiserar av Hägg för »en mycket snäv definition av placebobegreppet«, där psykiska och verbala stimuli inte räknas, utan endast placebo givet som en behandlingsliknande åtgärd. Kritiken är inte utan problem, eftersom den utgår ifrån att begreppet skall utvidgas med psykologiska faktorer i vårdmiljön, som ju sällan är under kontroll i placebolitteraturen. Kienle skulle t ex inte ha kunnat analysera Beechers artikel ur den synvinkeln eftersom sådana faktorer inte redovisats i tidiga placebostudier.

I SBU-boken om patient-läkarrelationen [7] står de psykologiska faktorerna i förgrunden, och de är självklart av stort intresse. Frågan huruvida de bör inlemmas i placebobegreppet har flera sidor. Inte minst är det tveksamt om man vinner något med att läkare och vårdpersonal som ger ett gott bemötande börjar betrakta sig som förmedlare av placeboeffekter. Jag kan tycka att det psykologiska omhändertagandet är värt en mer självständig ställning än så. Placebobegreppet torde också bli än mer svårhanterligt om man utökar det med allt som har specifik stödjande verkan i vård situationen.

Hägg diskuterar den stora variabiliteten hos placeboeffekterna, vilket är ett konstruktivt avsteg från traditionen. Han väljer också ut tillstånd som är lättare att studera: graviditetsillamående, astma, kontaktdermatit, smärta, hypertoni. Här ligger det ju av olika skäl nära till hands att suggestionseffekter åtminstone kortsiktigt kan spela en roll. En empiriskt grundad placeboforskning måste rimligen ta sig an mer löstlösta problem i första hand och akta sig för de svepande generaliseringar som präglat området under ett halvsekel.

Karin Johannisson's kapitel i boken håller en välgörande distans till de internt medicinska motiv som präglat synen på placebo. Hon talar bl a om »en arrogant slutsats: fram till 1900-talet handlade i stort sett all läkemedelsbehandling om placebo, eftersom verkningsfulla medel saknades«. Med frasen »bara placebo« har mycket avvisats som den etablerade medicinen inte vill acceptera.

Frågan om varaktigheten av placeboeffekter berörs på några ställen. Boström anser att litteraturen tyder på att den varierar mellan någon vecka och flera månader, men ger inga referenser. Hägg citerar en studie av angina pectoris, där effekt eventuellt förelåg under en observationstid på sex månader. Marie Åsberg diskuterar möjligheten av placeboeffekt vid långvarig depressionsbehandling och anser sig inte ha funnit något säkert svar. Tillgången på evidens av ens det enklaste slag är tydligen mycket liten när det gäller durationsfrågan. Detta måste hållas i minnet när placeboeffekter påstås förklara tillfrisknanden från långvarig sjukdom.

Kommentarer i Läkartidningen

Placebodebatten har under det senaste året kommenterats i Läkartidningen vid två tillfällen, först av Jan-Otto Ottosson [9] och senast i januari 2002 av Erik Hägg och Harry Boström [10]. Ottosson's argumentation utgår ifrån att jag i en artikel ifråga-

satt att placeboeffekter alls existerar, vilket är en missuppfattning [11]. Även Hägg och Boström förutsätter att placebobegreppet skall försvaras mot rena förnekare, och här är det artikeln av Hróbjartsson och Gøtzsche [5] som står i fokus, eller snarare kommentarerna till den i olika medier världen runt [12].

Ottosson hänvisar till den enligt honom övertygande dokumentation som finns i boken om patient-läkarrelationen [7] och använder valda delar av detta material i sin artikel. Hägg och Boström sammanfattar på motsvarande sätt litteratur som de använt i boken »Placebo« [8]. Dessutom för de kritiska resonemang kring den aktuella danska undersökningen, som givetvis har begränsningar men dock är det hittills största systematiska försöket att påvisa »sanna« placeboeffekter. Hróbjartsson och Gøtzsche kunde inte alls finna sådana tydliga effekter som motsvarar den gängse uppfattningen om placebo's »potens«. Hägg och Boström radar upp invändningar så att läsaren kan få intrycket att detta inte behöver betyda särskilt mycket, vilket faktiskt är missvisande.

Överdrivna föreställningar

Vi vet idag att våra föreställningar om placeboeffekter har varit överdrivna från första början, men reträtten från tidigare positioner är inte alltid öppen och oreserverad. En omvärdering är nödvändig; detta ändras inte av att PET-bilder ser ut att konkret påvisa en verkan av placebo på hjärnan i vissa situationer [13, 14]. Att ett fenomen existerar betyder ju inte att det också har alla egenskaper som traditionen tillskrivit det.

Drömmen om en panacé, medlet som botar allt, har funnits i medicinen sedan antiken. I läran om placebo lever den på sätt och vis kvar. Nu måste tiden vara inne för nykter analys, där alla påståenden om placebo prövas strängt mot den empiriska verkligheten. Placeboeffekten existerar, men vi behöver veta hur mycket utöver detta magra konstaterande som vilar på solid evidens. Tack vare kritiker som Kienle, Hróbjartsson och Gøtzsche har äntligen arbetet börjat!

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Kienle GS. Der sogenannte Placeboeffekt. Stuttgart: Schattauer; 1995.
2. Beecher HK. The powerful placebo. JAMA 1955;159:1602-6.
3. Kienle GS, Kiene H. Placebo effect and placebo concept: a critical methodological and conceptual analysis of reports on the magnitude of the placebo effect. Altern Ther Health Med 1996;2:39-54.
4. Kienle GS, Kiene H. The powerful placebo effect: fact or fiction? J Clin Epidemiol 1997;50:1311-8.
5. Hróbjartsson A, Gøtzsche PC. Is the placebo powerless? – An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 2001;344:1594-602.
6. Bailar JC III. The powerful placebo and the wizard of Oz. N Engl J Med 2001;344:1630-2.
7. Ottosson JO, red. Patient-läkarrelationen: läkekonst på vetenskaplig grund. Stockholm: Natur och Kultur/SBU; 1999.
8. Boström H, Dahlgren H, red. Placebo. Stockholm: Liber/SBU; 2000.
9. Ottosson JO. Placeboeffekten behöver inte betvivlas! Läkartidningen 2001;98:2434-6.
10. Hägg E, Boström H. Placebo – något att tro på? Läkartidningen 2002;99:261-2.
11. Dalén P. Rutinmässig användning av placebo ifrågasätts. Tydligt ställningstagande i Helsingforsdeklarationen. Läkartidningen 2001; 98:16-8.
12. Gibbs WW. All in the mind: Fact or artifact? The placebo effect may be a little of both. Sci Am 2001;285(4):10.
13. Kahl U. Placebo – myt eller vetenskap? Läkartidningen 2001; 98: 5001.
14. Petrovic P, Kalso E, Petersson KM, Ingvar M. Placebo and opioid analgesia. Imaging a shared neuronal network. Science 2002; 295: 1737-40.



**EVIDENSBASERADE STUDIER
- REFERAT OCH
KOMMENTARER**

ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Antibiotikas effekt på symtom och komplikationer vid halsinfektion

Kommentar till en metaanalys från Cochrane Collaboration

Det rekommenderas ofta att patienter med halsinfektion orsakad av grupp A-streptokocker (GAS) bör behandlas med antibiotika. Syftet är att lindra symtomen, förkorta sjukdomsdurationen, reducera förekomsten av variga komplikationer (t ex peritonsillitabscess), att reducera förekomsten av icke-variga komplikationer (reumatisk feber och akut glomerulonefrit) samt att begränsa smittspridning och därmed avbryta smittkedjor. Denna översikt kommenterar en systematisk litteraturoversikt från Cochrane Collaboration avseende antibiotikas inverkan på symtom och komplikationer vid halsinfektion [1].

Metaanalysens material och metoder

Metaanalysen är baserad på en systematisk genomgång av litteraturen sedan 1945 och på elektronisk sökning i Medline från 1966 (sökord: pharyngitis, sore throat och tonsillitis).

De publicerade arbetena inkluderades i metaanalysen om de omfattade en prövning av ett antibiotikum i jämförelse med placebo och det fanns ett mått på effekten på den akuta sjukdomens symtom eller efterföljande uppträdande av variga eller icke-variga komplikationer. Effekten på minst ett av följande utfall skulle vara mätt: 1. Incidensen av reumatisk feber inom två månader, akut glomerulonefrit inom en månad, akut otitis media, akut sinusit eller peritonsillitabscess. 2. Angivande av följande

Författare

STEEN HOFFMAN

överläkare, Neisseriaafsnittet, Statens Serum Institut, Köpenhamn, Danmark

E-post: hof@ssi.dk

HANS JØRN KOLMOS

överläkare, dr med, klinisk mikrobiologisk afd, H:S Hvidovre Hospital, Danmark.

Denna artikel är en kritisk genomgång av en Cochrane-översikt. Manuskriptet har genomgått extern referentbedömning. Artikeln är tidigare publicerad i Ugeskr Læger 1998; 160: 3373-5.

symtom: ont i halsen, huvudvärk eller feber.

Metaanalysens resultat

Arton arbeten omfattande 9 189 fall av halsinfektion ingick i metaanalysen; elva är publicerade under 1950-talet och sju i perioden 1977–1992. Elva arbeten exkluderades, några på grund av att det saknades mått för effektvariabler, andra för att det inte förelåg en placebobebehandlad kontrollgrupp, men t ex jämförelse mellan två olika antibiotika eller jämförelse av olika behandlingduration. Metaanalysens resultat är uttryckta som relativ risk, estimerad med Petos odds ratio med 95 procents säkerhetsgränser, Tabell I.

Icke-variga komplikationer: Incidensen av akut glomerulonefrit var endast angiven i sex arbeten. Där fanns två fall och bägge inträffade hos obehandlade patienter. Den relativa risken beräknades till 0,07, men med så vida konfidensintervall (0,00–1,32) att man inte med detta som grund kan hävda att akut glomerulonefrit kan förebyggas med antibiotika. Däremot reducerades risken för att utveckla reumatisk feber till omkring en tredjedel av risken hos obehandlade.

Variga komplikationer: Bland de antibiotikabehandlade patienterna var risken för att utveckla akut otitis media reducerad till cirka en fjärdedel av risken bland de obehandlade. Risken för akut sinusit och peritonsillitabscess reducerades till cirka en tredjedel respektive cirka en sjättedel.

Symtom: Bland de behandlade patienterna reducerades förekomsten av ont i halsen, huvudvärk och feber tre dagar efter det att patienten inkluderats i respektive undersökning till cirka hälften av förekomsten hos de obehandlade. Både behandlade och obehandlade patienter var i 90 procent symtomfria en vecka efter symtomdebuten. Det beräknas att antibiotika förkortade symtomens duration med genomsnittligt åtta timmar.

Metaanalysens diskussion och konklusioner

Författarna lägger i diskussionen stor vikt vid en jämförelse mellan de relativa och de absoluta fördelarna med

Sammanfattat

- Arton studier och drygt 9 000 patienter inkluderades i en systematisk översikt från Cochrane Collaboration avseende effekten av antibiotikabehandling vid halsinfektion.
- En kritisk analys av översikten visar att i endast fem av studierna var det ett specificerat krav att patienterna skulle ha verifierade grupp A-streptokocker (GAS) och att resultaten alltså belyser nyttan av antibiotika vid ospecifik halsinfektion, inte tonsillit.
- Jämfört med placebo reducerade antibiotikabehandling risk för reumatisk feber till en tredjedel, men antalet fall med glomerulonefrit var för litet för en säker analys.
- Risken för sinusit respektive peritonsillit reducerades till en tredjedel respektive en sjättedel av risken i den placebobebehandlade gruppen.
- Antibiotika förkortade symtomens varaktighet med cirka åtta timmar.
- Författarna argumenterar för att antibiotikabehandling vid GAS begränsar smittspridning och virulensökning och varnar för att använda den aktuella översikten som stöd för okritiskt undanhållande av antibiotika vid halsinfektioner.

antibiotikabehandling. De beräknade odds ratio-måtten är relativa, genom att de uttrycker t ex en halvering av en risk, medan en förståelse för de absoluta förhållandena uppnås genom att man ser på hur vanliga de faktiska komplikationerna är.

Exempelvis är den erhållna relativa

Tabell I. Effekt av antibiotika på komplikationer och symtom vid halsinfektion.

Komplikation eller symtom	Antal studier	Förekomst bland antibiotika-behandlade (n/N)	Förekomst bland placebo-behandlade (n/N)	OR	(95 procent säkerhetsgränser)
Akut glomerulonefrit	6	0/1 625	2/1 459	0,07	(0,00-1,32)
Reumatisk feber	10	28/4 208	87/3 532	0,28	(0,19-0,40)
Akut otitis media	7	11/1 305	27/845	0,23	(0,12-0,45)
Akut sinuit	3	3/466	4/202	0,33	(0,07-1,64)
Peritonsillar abscess	2	1/290	17/311	0,17	(0,07-0,44)
Ont i halsen på tredje dagen	8	449/1 113	526/851	0,39	(0,32-0,47)
Huvudvärk på tredje dagen	2	115/434	110/248	0,40	(0,28-0,56)
Feber på tredje dagen	6	83/679	119/601	0,56	(0,41-0,76)

Talen kommer från [1]

n/N = antal med anförd komplikation eller symtom/antal undersökta totalt
OR = odds ratio.

riskreduktionen av reumatisk feber mycket stor, men i länder där reumatisk feber bland obehandlade är ovanlig blir den absoluta risken minimal. Reumatisk feber är inte ovanlig i befolkningar med dåliga socioekonomiska villkor; metaanalysens författare anför som exempel aboriginerna i Australien och fortsätter: »Antibiotics may be justified to reduce the complications of acute rheumatic fever in these settings.»

Med hänvisning till en tidigare litteraturstudie utförd av metaanalysens försteförfattare [2] hävdas att 30 barn eller 145 vuxna måste behandlas med antibiotika för att förebygga ett fall av akut otitis media i dessa åldersgrupper.

Det nämns i diskussionen att det hade varit önskvärt att även innefatta biverkningar och andra oväntade effekter av antibiotika, men detta är dessvärre inte gjort. Vidare påpekas att det inte har varit möjligt att analysera antibiotikas goda effekt stratifierat efter symtomens svårighetsgrad, men att en sådan stratifierad analys kanske i högre grad skulle ha visat en fördel med antibiotikabehandling.

I metaanalysens konklusion framhålls att i socioekonomiskt sämre ställda befolkningar kan antibiotika vara indicerat för att reducera antalet fall av reumatisk feber. I andra befolkningsgrupper, där denna komplikation är ovanlig, anser metaanalysens författare att man bör väga den lätt symtomreducerande effekten av antibiotika mot de risker som är förbundna med antibiotikabehandling. Författarna anför inga ytterligare rekommendationer.

Vår kommentar

En genomgående och grundläggande svaghet i denna metaanalys är att det läggs alltför ringa vikt – om över huvud taget någon – vid att skilja mellan förhållandena vid streptokockbetingade och virala fall. Detta gäller såväl litteraturanalysen som – i synnerhet – den efterföljande diskussionen. I konklusio-

nerna ingår över huvud taget inga överväganden om att den kliniska skillnaden mellan streptokockbetingade och virala fall är opålitlig och att undersökning av en svalgodling därför krävs, t ex genom direkt påvisande av GAS-antigen. Metaanalysen tycks utgå från samhällen där mikrobiologiska undersökningar av svalgodlingar knappast är tillgängliga, vare sig de sker på laboratorier eller decentraliserat i primärsektorn.

I bara 5 av de 18 studierna var det ett specificerat krav att patienterna skulle ha odlingspåvisad GAS. Man har ändå inkluderat ett arbete i vilket patienter med GAS specifikt blev uteslutna (en undersökning av amoxicillin, sulfametoxazol–trimetoprim och placebo vid förmodade virala luftvägsinfektioner hos barn). Det framgår indirekt av metaanalysen att man inte har försökt att exkludera de patienter som i förhållande till originalartikelns upplysningar inte hade GAS. Vid läsning av originalartiklarna noterar vi att en sådan exklusion skulle ha varit möjlig i en lång rad fall, och att i åtskilliga artiklar framhävs ändå att penicillinbehandling hade en slående effekt på symtomen hos de streptokockinficerade, men inte hos patienter utan GAS. Man har också tagit med arbeten där det inte ingick en gruppbestämning av de betahemolytiska streptokocker man funnit. Det okritiska urvalet av patienter minskar den eventuella effekten av penicillin på GAS-tonsillit.

Det är inte något nytt att penicillin kan förebygga reumatisk feber. Att penicillin skulle kunna förebygga akut glomerulonefrit har man betvivlat i många år [3], så inte heller på denna punkt bidrar metaanalysen med nya rön. Både akut sinuit och akut otitis media kan i hög grad, och peritonsillitabscess i någon mån, skyllas på andra bakterier än GAS, och ett förebyggande av dessa sjukdomar kan bara i ringa omfattning uppnås med penicillinbehandling av streptokocktonsillit. Det egentli-

gen intressanta i metaanalysen blir därför dess belysning av effekten på symtom.

Vid jämförelse av metaanalysens resultatpresentation (Tabell I) med de åtta originalartiklarna, som har beskrivit hur symtomet »ont i halsen» har fortskridit, har vi funnit att metaanalysens författare konsekvent har undvikit att skilja mellan patienter med och utan streptokocker. Dessa tal har helt enkelt slagits ihop. I tre av artiklarna finns inte tabellariska resultatsammanställningar, och metaanalysens tal måste här antas vara baserade på avläsning av kurvor. Bara tre av studierna rörde penicillin kontra placebo (eller bara acetylsalicylsyra), tre tog upp penicillin samt sulfa- och tetracyclinpreparat och de sista två studierna rörde sulfapreparat.

Vi finner det felaktigt att resultat som uppnås med preparat som inte är likvärdiga (som sulfa- och tetracyclinpreparat) slås ihop och behandlas gemensamt tillsammans med penicillin i metaanalysen. Man har således inte presterat en sund analys av relevanta antibiotikas inverkan på symtom vid streptokockorsakad halsinfektion. Man har heller inte inkluderat en relativt ny amerikansk studie, där det redan under det första dygnets behandling var en slående effekt av penicillin på symtomen av streptokocktonsillit hos barn [4].

Vi medger att metaanalysen, enligt titeln, behandlar ont i halsen i bred mening, alltså utan att man tar hänsyn till den mikrobiella etiologin. Vi fruktar dock att korfattade citeringar av metaanalysens resultat och konklusioner kan förleda de flesta läsare till att tro att det rör sig om penicillins effekt på GAS-tonsillit och inte bara på en dåligt definerad blandning av virala och bakteriella infektioner. Utifrån detta betraktande förefaller oss metaanalysen att vara starkt vilseledande.

Det är i metaanalysen inte taget ställning till att antibiotikabehandling av den enskilda patienten kan bidra till att avbryta smittkedjor och därmed förhindra, att streptokockernas virulens ökas. GAS har i sin cellvägg flera ytproteiner, av vilka M-protein är ett typspecifikt antigen och en virulensfaktor med antifagocytära egenskaper. Undersökningar från bl a militärförläggningar i USA har visat att GAS-stammar som isolerats under pågående epidemier innehåller mer M-protein än stammar som isolerades under endemiska förhållanden. Vid in vitro-försök har man iakttagit, att om en streptokockstam hålls levande genom att man ofta startar nya odlingar från en platta till en annan bortfaller gradvis produktionen av M-protein, men genom efterföljande passager in vivo, t ex från en mus till en annan, börjar stammen åter att producera

M-protein [5]. I familjeundersökningar från USA under 1950-talet fann man att 25–50 procent av syskon till ett barn med streptokocktonsillit fick motsvarande sjukdom och att denna andel var större ju senare antibiotikabehandling av indexbarnet sattes igång [6].

Även om vi är eniga om att den absoluta effekten av penicillin på risken att utveckla reumatisk feber är liten i ett modernt samhälle, som t ex Danmark, finner vi det viktigt att man är uppmärksam på de plötsliga anhopningarna av fall av reumatisk feber som förekom i mitten eller slutet av 1980-talet i USA [7] och i en rad europeiska länder. Helt i motsättning till tidigare förhållanden drabbades patienter från den vita medelklassen, som inte var fattiga och som hade god tillgång till läkare. Nära en tredjedel av patienterna med reumatisk feber hade inte fått antibiotika mot den inledande halsinfektionen. Det var en markant ökad förekomst av mukoida (dvs slembildande) GAS-stammar; förmågan att bilda mukoida kolonier på blodagerplattor är relaterad till stammarnas virulens. Under 1990-talet har man i både Danmark [8], USA [9] och en rad andra länder observerat en ökad förekomst av aggressiva nekrotiserande mjukdelsinfektioner förorsakade av GAS, inte minst hos yngre och i övrigt friska personer, och i flera av dessa fall har det varit tal om mukoida stammar.

Vår konklusion

Resultaten i denna metaanalys har inte ändrat vår uppfattning att GAS är virulenta och potentiellt livsfarliga bakterier, som bör behandlas med penicillin, när de påvisas hos patienter med relevanta kliniska symtom, t ex halsinfektion. Även om effekten på symtomens duration kanske inte är lika dramatisk hos alla patienter, och även om sjukdomen som regel är självlimiterande, har antibiotikabehandling en rad andra potentiellt gynnsamma effekter i denna situation, främst möjlighet att begränsa smittspridning med därav följande möjliga virulensökning.

*

Denna artikel kommenterar en översikt från the Cochrane Collaboration från 1997. I enlighet med Cochranes principer uppdateras översikter när nya studier tillkommer. I september 1999 publicerades en ny version av översikten som nu inkluderar totalt 10 484 patienter. Slutsatserna förändrades dock inte.

Referenser

1. Del Mar CB, Glasziou PP. Antibiotic for the symptoms and complications of sore throat. I: Douglas R, Bridges-Webb C, Glasziou P, Lozano J, Steinhoff M, Wang E, eds. Acute respiratory infections module of the Coch-

rane Database of Systematic Reviews (updated 01 September 1997). Available in the Cochrane Library (database on disk and CDROM). The Cochrane Collaboration; Issue 4. Oxford: Update Software; 1997. Updated quarterly.

2. Del Mar C. Managing sore throat: a literature review. II. Do antibiotics confer benefit? *Med J Aust* 1992; 156: 644-9.
3. Weinstein L, Le Frock J. Does antimicrobial therapy of streptococcal pharyngitis or pyoderma alter the risk of glomerulonephritis? *J Infect Dis* 1971; 124: 229-31.
4. Randolph MF, Gerber MA, DeMeo KK, Wright L. Effect of antibiotic therapy on the clinical course of streptococcal pharyngitis. *J Pediatr* 1985; 106: 870-5.
5. Rothbard S, Watson RF. Variation occurring in group A streptococci during human infection. Progressive loss of M substance correlated with increasing susceptibility to bacteriostasis. *J Exp Med* 1948; 87: 521-33.
6. Breese BB, Disney FA. Factors influencing the spread of beta hemolytic streptococcal infections within the family group. *Pediatrics* 1956; 17: 834-8.
7. Bisno AL. The resurgence of acute rheumatic fever in the United States. *Annu Rev Med* 1990; 41: 319-29.
8. Kristensen B, Egelund B, Meyer M, Henriksen J, Jepsen OB, Sievers CJ et al. Et udbrud af infektioner med *Streptococcus pyogenes* på institutioner for psykisk udviklingshæmmede i Storkøbenhavn 1995. *Ugeskr Læger* 1996; 158: 1679-82.
9. Stevens DL. Streptococcal toxic-shock syndrome: spectrum of disease, pathogenesis, and new concepts of treatment. *Emerg Infect Dis* 1995; 1: 69-78.

Särtryck

LÄKARTIDNINGEN

Palliativ medicin är vetenskapen som den palliativa vården grundar sig på, dvs utbildningen, utvecklingen och forskningen.

I Läkartidningen har under 1999 och 2000 publicerats 15 artiklar som belyser palliativ vård från olika sidor. Dessa har samlats i ett särtryck som kan beställas med kupongen nedan.

Priset är 80 kronor.

Palliativ medicin



Beställer härmed.....ex av "Palliativ medicin"

.....namn

.....adress

.....postnummer

.....postadress

Insändes till LÄKARTIDNINGEN
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se

Besök av facilitator och effekten på klinisk praxis och patientnytta

En Cochrane-översikt om "outreach visits"

I ökande omfattning utvecklas och implementeras strategier för optimering av kvaliteten i det danska hälso- och sjukhusväsendet. I allmänpraktik prövas systematisk efterutbildning, och utvecklingen av den kliniska vägledningen intensifieras. Audit Projekt Odense (APO) erbjuder registreringsprojekt med feedback och kursaktivitet. I Århus Amt har man i ett samarbete mellan allmänpraktiker och sjukhusen utvecklat State of the art-dokument.

Från sjukhussektorn finns erbjudande om efterutbildning, utveckling av klinisk vägledning samt ökad centralisering av vissa behandlingar med bättre kvalitet som mål. Endast i ringa omfattning används besök av en speciellt utbildad person, en så kallad facilitator. Det har skett i samband med kvalitetsssäkring inom t ex klinisk kemi, mikrobiologi och farmakologi. Bruket av facilitatorer är mera känt i utlandet, och speciellt i Storbritannien tycks man ha goda erfarenheter av att använda facilitatorer [1].

Det stora utbudet av olika strategier för kvalitetsförbättring gör det naturligt att undersöka om metoderna har effekt och vilka metoder som i så fall är mest effektiva, mätta med relevanta effektmått.

Syftet med denna artikel är att gå igenom en nyligen publicerad Cochrane-översikt [2] om effekten av »outreach visits» (besök av facilitatorer) mätt som förändring i klinisk praxis och patientnytta. Effekten av besök av facilitatorer jämförs med andra metoder för kvalitetsutveckling.

Resultat

I genomgången definierades begreppet »outreach visits» som ett personligt

Författare

PETER VEDSTED

leg läkare, Ph D studerande

JØRGEN LOUS

lektor, dr med, båda vid Forskningsenheden og Institut for almen medicin, Aarhus Universitet, Venneyst Boulevard 6, DK-8000 Århus.

Artikeln är tidigare publicerad i Ugeskr Læger 1998; 160: 7243-5.

besök av en speciellt utbildad person hos en sjukvårdsutbildad person på dennes arbetsplats. Detta är också känt som »academic detailing» och »university-based educational detailing». På danska har vi valt begreppet »facilitator».

Cochrane-översikten är genomförd systematiskt efter krav som är uppställda av The Cochrane Collaboration [3] av en grupp inom Cochrane Collaboration on Effective Professional Practice (CCEPP). Med sökning i Medline och Cinahl samt en genomgång av referenslistor från funna artiklar fann man 24 studier som uppfyllde inklusionskriterierna:

- Randomiserad eller kvasirandomiserad allokering till interventionsgrupp(er).
- Omfattar sjukvårdspersonal (exkluderande studenter).
- Personligt besök av en speciellt utbildad person hos en behandlare i hans/hennes egna lokaler med eller utan kompletterande material eller aktiviteter.
- Objektiva mått på klinisk praxis och/eller nytta för patienterna.

Varje studie bedömdes av två av varandra oberoende gruppdeltagare avseende kvaliteten på studien utifrån sex kriterier. Därtill noterades metod, material, intervention och effekt. Sex studier exkluderades på grund av att det saknades randomisering, 18 studier ingick därmed.

Tabell I visar resultatet av genomgången. Studierna har grupperats utifrån vilka åtgärder som facilitatorns besök jämfördes med. Det är sammanlagt 19 jämförelser i tabellen. Det beror på att en studie jämförde dels besök av en facilitator i kombination med andra insatser med ingen intervention, dels jämförde besök av en facilitator med audit och feedback.

13 av de 18 studierna behandlade problemställningar kring ordination av läkemedel, t ex reducerad användning av benzodiazepiner, NSAID, tetracyklin och blodtransfusioner, ökad användning av lipidsänkande medicin och reducerade läkemedelskostnader. Tre studier handlade om profylaktiska åtgärder. Två studier fokuserade på förbättrad behandling av vanligt förekom-

Sammanfattat

- »Outreach visits» definieras som ett personligt besök av en speciellt utbildad person hos en sjukvårdsutbildad person på dennes arbetsplats. På danska används begreppet »facilitator».
- Syftet med denna artikel är att gå igenom en nyligen publicerad Cochrane-översikt om effekten av besök av facilitatorer mätt som förändring i klinisk praxis och patientnytta. Effekten av besöken jämförs med andra metoder för kvalitetsutveckling.
- Genomgången visar att man helt saknade studier fokuserade på användning av »facilitatorer».
- Dock kan konkluderas att besök av en facilitator är en effektiv metod för att ändra ordinationsvanor. Effekten på patientnyttan var inte överbevisande.
- Innan det föreligger bevis för att metoden också har betydelse på patientnivå bör den bara implementeras på försök med inbyggd monitorering.

mande sjukdomar i allmän praktik (astma, diabetes, otitis media, hypertension, akut bronkit samt ångest och depression).

Alla studier använde effektmått, som mätte läkarnas kliniska praxis. Tre studier använde indikatorer som mätte annan sjukvårdspersonals kliniska arbete. Bara en studie mätte effekten på patientnyttan efter besök av en facilitator. Det var en studie som fokuserade på användningen av psykofarmaka och antidepressiva. Här fann man en signifikant bättre behandling av depression värderad av läkare, men inte någon mätbar ef-

Tabell I. Resultaten i 18 randomiserade undersökningar om effekten av besök av facilitator. Sammanlagt 19 jämförelser med andra interventioner eller ingen intervention.

Jämförelse	Antal	Material	Indikator för måluppfyllelse	Effekt, procent
Besök av facilitator /ingen intervention	3	Allmänpraktiserande läkare	Färre ordinationer Fler ordinationer	14–24 38
Besök av facilitator + reminder, audit, feedback eller skriftligt material/ingen intervention	13	Allmänpraktiserande läkare, sköterskor och andra	Förbättrad ordination Minskat bruk av blodtransfusion Fler förebyggande råd	1–45 46 9–27
Besök av facilitator /audit och feedback	1	Sjukhusläkare	Sänkta ordinationsutgifter	11
Besök av facilitator + patientrelaterad undervisning/besök av facilitator	1	Läkare och sköterskor	Bättre ordination	9
Besök av facilitator /effekt över tid	1	Läkare och sköterskor	Bättre ordination	Ingen skillnad

fekt med hänsyn till mental funktion, minnesfunktion, ångest, beteende och sömn.

Mätt utifrån utvalda indikatorer visade alla studierna att facilitatorsbesöken hade en positiv effekt på det kliniska arbetet. Samtidigt visade sig besök av en facilitator vara mer effektiva än de övriga metoderna som den jämfördes med; audit (medicinsk revision), reminder, feedback, skriftligt material (Tabell I). Studierna visade dessutom en ökad effekt om facilitatorns besök kombinerades med andra metoder.

Kommentarer

Vid genomgången av den inkluderade och exkluderade litteraturen fann vi det anmärkningsvärt att man helt saknade studier fokuserade på användning av »facilitatorer», som är beteckningen för »out-reach visits» i Storbritannien. Detta system täcker allmänmedicinska mottagningar i hela Storbritannien med omkring 500 anställda. I Storbritannien är en facilitator en speciellt utbildad sjukvårdsperson, oftast sjuksköterska, som med jämna mellanrum besöker en grupp läkarmottagningar för att introducera och implementera god klinisk praxis i läkar- och sjuksköterskegruppen [1].

Det visade sig att den sökstrategi man hade använt i Cochrane-översikten inte hittade studier om facilitatorsystemet. I Cochrane-gruppen är man medveten om detta och planlägger en ny genomgång av dessa studier [Thomson MA, pers medd].

Konklusion och perspektiv

Mätt utifrån utvalda indikatorer hade besök av en facilitator hos en läkare en positiv effekt på det kliniska arbetet. Effekten kunde jämföras med eller vara större än den som uppmättes vid användning av andra metoder. Antalet randomiserade studier är dock begränsat och rör sig övervägande om ordina-

tion av medicin. På många områden är det således behov av mera forskning om effekten av besök av en facilitator, och effekten bör jämföras med andra metoder för kvalitetsutveckling. Dessutom behövs en värdering av facilitatorsbesökens kostnadseffektivitet.

Det kan dock konkluderas att besök av en facilitator är en effektiv metod för att ändra ordinationsvanor. Bara en studie värderade facilitatorsbesökens effekt på patientnyttan. Effekten var inte överbevisande, vilket klart antyder att innan det föreligger bevis för att metoden också har betydelse på patientnivå, bör den bara implementeras på försök med inbyggd monitorering.

En ny Cochrane-genomgång av studier av facilitatorsystemet är planlagd och bör emotses med stort intresse.

Referenser

1. Vedsted P, Lous J. Kvalitetsudvikling ved besøg af praksiskonkulent i almen praksis. Erfaringer fra England og Holland. *Practicus* 1998; 22: 90-2.
2. Thomson MA, Oxman AD, Davis DA, Haynes RB, Freemantle N, Harvey EL. Out-reach visits to improve health professional practice and health care outcomes. In: Bero I, Grilli R, Grimshaw J, Oxman A, eds. Collaboration on effective professional practice module of the The Cochrane Database of Systematic Reviews, (updated 1 December 1997). Available in the Cochrane Library (database on disk and CDROM). The Cochrane Collaboration; issue 1. Oxford: Update Software, 1998.
3. Mulrow CD, Oxman A, eds. How to conduct a Cochrane Systematic Review. San Antonio, Texas: The Cochrane Collaboration, 1997.

Palliativ medicin är vetenskapen som den palliativa vården grundar sig på, dvs utbildningen, utvecklingen och forskningen.

I Läkartidningen har under 1999 och 2000 publicerats 15 artiklar som belyser palliativ vård från olika sidor. Dessa har samlats i ett särtryck som kan beställas med kupongen nedan.

Priset är 80 kronor.

Palliativ medicin



Beställer härmed.....ex av "Palliativ medicin"

.....namn

.....adress

.....postnummer

.....postadress

Insändes till **LÄKARTIDNINGEN**
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se

SBU-rapporten om ont i rygg och nacke

Ambitiös genomgång av ett stort problemområde

Se artikeln
på sidan 4974
i detta nummer.

Smärtor i rygg och nacke är vanligt förekommande i den allmänna befolkningen, något som har visats i flera epidemiologiska studier. Nacksymtomen har sannolikt ökat under senare år, medan andelen med ländryggssmärter legat kvar på ungefär samma nivå [1]. Livstidsprevalensen för ryggbesvär uppgår till över 80 procent, för nackbesvär till ca 50 procent, vilket ger en uppfattning om dimensionen på ett problem som ofta medför läkarkontakt.

Smärta lokaliserad till rörelseorganen är också en av de vanligaste orsakerna till besök i primärvården. Omkring 15 procent av besöken, alltså var sjunde patientkontakt, beror på denna typ av besvär [2]. Handläggningen i primärvården präglas och utformas utifrån denna organisations förutsättningar och organisation.

Det finns stora utvecklingsmöjligheter

Utvecklingsmöjligheter finns, och sådana behöver stimuleras och understödjas [3]. Formaliserad konsultverksamhet med specialist från sjukhusvård kan effektivisera handläggningen av patienter med besvär från rörelseorganen samt ge samhällsekonomiska vinster i form av minskat sjukskrivningsbe-

hov [4]. Det synes självklart att diagnostiska och terapeutiska åtgärder vad gäller smärta i rygg och nacke skall vara väl underbyggda ur både individens och samhällets perspektiv.

Kunskapsmängden har vuxit betydligt

Evidensbaserad medicin har under 1990-talet med rätta vunnit stora insteg i klinisk praxis inom de flesta specialiteter. Utmönstring av dåligt underbyggda diagnostiska och terapeutiska metoder är viktig, främst från den enskilde patientens synpunkt, men också i ett samhällsperspektiv i en tid av minskande gemensamma resurser. Redan 1991 publicerade Statens beredning för medicinsk utvärdering (SBU) en uppmärksam rapport om ont i ryggen [5], men sedan dess har kunskapsmängden vuxit betydligt.

Ur de ca 25 000 vetenskapliga artiklar som idag behandlar smärta i rygg och nacke har en expertgrupp vaskat fram 2 000 studier med den starkaste evidensen, och under fyra år bearbetat och sammanställt sina resultat.

I detta nummer av Läkartidningen presenteras en sammanfattning av SBUs nya evidensbaserade kunskaps-sammanställning »Ont i ryggen, ont i nacken» [6]. I rapporten poängteras att »ryggsmärter och dess konsekvenser är tillsammans inte ett isolerat kroppsligt fenomen». Betydelsen av en biopsykosocial modell för uppkomst och underhåll av smärttillståndet, som diskuteras redan i den första rapporten, understryks nu med än större tyngd.

Multidisciplinärt arbetssätt med vårdteam

Adekvat handläggning förutsätter då ofta ett multidisciplinärt arbetssätt med inslag av läkekonst. Först då kan effekter förväntas på såväl individnivå som samhällsnivå. Som tidigare nämnts handläggs många av dessa patienter i primärvården; i SBU-rapporten flaggas också för flera principer som torde kun-

na anammas av en primärvård i utveckling. Ett multidisciplinärt arbetssätt där ett team bestående av läkare, sjukgymnast och kurator (eller annan psykosocial kompetens), eventuellt kompletterad med ytterligare yrkeskategorier, gemensamt effektivt kan bedöma och behandla individer med långvarig och sammansatt problematik bör prövas.

Nyttan av »röda» flaggor, som signaler på biomedicinska avvikelser, och dito »gula», som flaggar för psykosociala riskfaktorer baserade på vetenskaplig evidens, bör vara stor i de tidiga patientkontaktarna. Helhetssyn, läkekonst och aktiv exspektans, ord som kännetecknar en välfungerande primärvård, framstår i SBU-rapporten som viktiga komponenter i handläggningen av rygg- och nacksmärter. Med rätta ställs primärvården i centrum för de insatser som oftast är tillräckliga. Det är betydelsefullt att mötets och iatrogena faktorer (t ex överutredning) påverkan på prognosen betonas.

Både sjukfrånvaro och förtidspensionering har ökat

Rörelseorganens sjukdomar svarar för 45 procent av antalet förtidspensioner i Sverige, varav ländryggbesvär står för 22 procent (1995). Totalt fanns i december 1999 ca 425 000 förtidspensionärer i Sverige [7]. Andelen förtidspensionärer har ökat de senaste decennierna. År 1980 var 5,6 procent av Sveriges befolkning i åldrarna 16–64 år förtidspensionärer, en andel som ökade till 6,7 procent 1990 och till 7,6 procent år 1998 [8].

Sjukfrånvaron har också dramatiskt ökat i Sverige sedan första kvartalet 1998, framför allt bland offentligtanställda [9]. Antalet pågående sjukfall ökade under 1998 med 32,9 procent för kvinnor och 26,5 procent för män, vilket skall jämföras med 4,4 procent respektive 1,4 procent under 1997. Bland tänkbara orsaker diskuteras attityder (sjukförsäkringen ses som frånvaroför-

Författare

H INGEMAR ANDERSSON

med dr, universitetslektor, Högskolan, Kristianstad; distriktsläkare, Vårdcentralen, Bromölla; överläkare, smärtrehabilitering, Sjukhuset, Kristianstad

E-post: ingemar.andersson@hv.hkr.se

IDO LEDEN

överläkare reumatologsektionen, medicinska kliniken, Sjukhuset, Kristianstad.

E-post: ido leden@skane.se

säkring) och det hårdnande arbetsmarknadsklimatet.

Systematisk värdering

Politiskt har alltid arbetslinjen hävdats i Sverige [10]. I SBU-rapporten framförs evidens för att aktivitet och inriktning på bibehållande av yrkesverksamhet bidrar till att minska risken för smärtbeteende och sjukrollsutveckling. Med hänsyn till utvecklingen av sjukförsäkringsdata har vi kanske inte fullt ut tagit till oss detta budskap.

Olika typer av konkreta åtgärder vad gäller rygg- och nacksmärtor är systematiskt värderade. Många fysikaliska behandlingsmetoder som fortfarande idag används flitigt utmönstras oftast på grund av svag eller ingen evidens i tillgängliga studier. Viktigast är dock de både positiva och negativa slutsatser där evidensen är stark. Vid långvarig ryggsmärta finns t ex stark evidens för positiva effekter av kognitiv beteendeterapi, vilket bör beaktas i olika program för ryggrehabilitering. Det diagnostiska avsnittet värderar förtjänstfullt evidensen för olika röntgenundersökningar av rygg och nacke.

Varning för medikalisering

Med rätta varnas för medikalisering och passiv, slentrianmässig handläggning. Tillika för att tidsbrist, insufficiens och bristande erfarenhet ofta leder till pseudoåtgärder, vilket ger falsk trygghet för både patient och sjukvård.

SBU-rapporten är en hjälp och ett sätt dels att motverka sådana tendenser, dels att utveckla och förfina handläggningen av ryggsmärtor.

Invändningar

Finns det då inga invändningar mot SBU-rapporten? Jo, men dessa handlar kanske mest om svagheter hos evidensbaserad medicin i allmänhet. Även om urvalet artiklar är stort och ambitiöst avspeglas måhända inte en klinisk verklighet. Är resultaten, som oftast hänför

sig till studier på utvalda grupper med aggregerade data, alltid applicerbara på en individuell, komplex situation? Effektiviteten av en viss åtgärd i vardagsituationen är kanske inte alltid möjlig att predicera, men samtidigt är gjorda studier det bästa vi har att bygga på.

En annan kritisk punkt är tillgång till och urval av studier. Flest och bäst studier finns inom ämnesområden som är mest »forskningsbara», och tillgången kan därmed påverka resultatet. I SBU-rapporten fastslås t ex att kurortsbehandling med hög evidens har en positiv effekt vid kronisk ländryggsmärta. Är kurortsbehandling verkligen en väldefinierad terapiform, jämförbar mellan olika studier, och kan kunskapen användas i praxis?

Bristen på adekvata studier inom flera områden har dock en viktig uppgift genom att peka framåt mot nya nödvändiga forskningsuppgifter inför framtiden. Så är t ex behandling av nacksmärtor betydligt sämre studerad än behandling av ryggsmärtor. Den nedslående evidensen för de flesta preventiva åtgärder borde också stimulera till intensivare forskningsinsatser på det förebyggande området.

Ett digert arbete

Sammanfattningsvis är vi dock mycket positiva till SBU:s digra arbete i två volymer på sammanlagt drygt 800 sidor. Genomgången av olika frågeställningar, värderingsprinciper och litteratururval är mestadels lättbegriplig och välgrundad, resultat och slutsatser är viktiga och intressanta för alla som möter rygg- och nackpatienter.

Nu återstår frågan om hur resultaten skall föras ut och förhoppningsvis få genomslag i den praktiska vardagen. Med hänsyn till de faktorer som visats påverka förekomsten av smärta i rygg och nacke räcker det dock inte att nå ut enbart till den medicinska professionen. Politiker, företagsledare och andra beslutsfattare borde begrunda många av rapportens slutsatser.

Referenser

1. Andersson HI. Chronic pain. Epidemiological studies in a general population [dissertation]. Lund: Lunds Universitet, 1998.
2. Andersson HI, Ejlertsson G, Scherstén B, Leden I. Musculoskeletal chronic pain in general practice. Studies of health care utilization in comparison with pain prevalence. *Scand J Prim Health Care* 1999; 17: 87-92.
3. Andersson HI, Leden I. Att möta individer med smärta från rörelseapparaten – tidiga överväganden i primärvård. *Allmänmedicin*. Under publ.
4. Leden I, Berglund P, Ejlertsson G, Henrysson I. Effektivare samverkan minskar behovet av sjukskrivning vid besvär i rörelseorganen. *Läkartidningen* 1991; 88: 2081-3.
5. Nachemson A. Ont i ryggen – orsaker, diagnostik och behandling. Stockholm: SBU, 1991.
6. SBU. Ont i ryggen – ont i nacken. En evidensbaserad kunskapssammanställning. Stockholm: SBU 2000. Rapport 2000: 145/1-2.
7. RFV. Statistikinformation IS-I 2000:003. Förtidspensioner utbetalade i december 1999. Stockholm: RFV 2000.
8. Edén L. Quality of life among disability pensioners with musculoskeletal disorders. Studies in a geographically defined population in Sweden [dissertation]. Lund: Lunds Universitet 2000. Meddelanden från Socialhögskolan 2000:6.
9. Mikaelsson B, Arrelöv B, Melander H, Normén G, Olsson B, Sjöberg I et al. Varför ökar sjukfrånvaron? *Läkartidningen* 1999; 96: 4160-2.
10. Lindqvist R, Marklund S. Forced to work and liberated from work. *Scand J Soc Welfare* 1995; 4: 224-37.

Smärtlindring för spädbarn

Översikt över randomiserade studier

Under 1990-talet introducerades EMLA (smärtstillande salva) och oral sockerlösning som smärtlindring vid blodprovstagning och andra smärre ingrepp på nyfödda barn. Metoderna har utvärderats med randomiserade, kontrollerade prövningar (randomized, controlled trials, RCT) som i sin tur har sammanfattats i systematiska översikter, främst från Cochranesamarbetet.

Arne Ohlsson, svensk pediatriker och professor i pediatrik i Toronto, och medarbetare står för flera randomiserade studier och Cochraneöversikter om neonatal smärta. De har nu sammanfattat egna och andras studier och systematiska översikter i ett bokkapitel [1]. Frågeställningen var: Vilken är den effektivaste och minst smärtsamma metoden för att erhålla ett blodprov för rutinmässig metabolisk screening (PKU-prov) av ett friskt, fullgånget, nyfött barn?

Den som inte har tillgång till boken [1] kan söka The Cochrane Database of Systematic Reviews via www.ronden.se med sökorden »OHLSSON and PAIN» och på så sätt ta del av en del av kunskapsunderlaget för kapitlet.

FOTO: FREDRIK FUNK, PRESSENS BILD.

METOD

Medline och Cochrane Library Databases genomsköptes i augusti 1998 med sökorden (Medical Subject Headings, MeSHs): [infant, newborn] AND [heel lance (prick, stick, stab)] AND [randomized controlled trial] OR [controlled trial] OR [random allocation] NOT [animal]. Personliga arkiv och referenslistorna i funna randomiserade studier genomsköptes av den kanadensiska gruppen. Två bedömare granskade oberoende av varandra de randomi-

»Vilken är den mest effektiva och minst smärtsamma metoden för att erhålla ett blodprov för rutinmässig metabolisk screening av ett fullgånget, nyfött barn?» Med den frågeställningen har metoder utvärderats i ett flertal randomiserade prövningar. Barnet på bilden har inget samband med artikeln.

serade studier man funnit och kvalitetsbedömde dem efter vedertagna mallar. De fall man bedömt olika diskuterades tills man nått konsensus. Sammanfattningen begränsades till studier av hög kvalitet.

RESULTAT Venpunktion

Venpunktion är mindre smärtsam, kräver färre stick och går snabbare än hälstick och är den metod som bör användas. EMLA tycks kunna minska smärtan vid venpunktion, men flera studier behövs. EMLA i enstaka doser på spädbarn medför inte biverkningar [1]. Enligt en fransk randomiserad stu-

Sammanfattat

- PKU-prov från nyfödda bör tas med venprov, inte med hälstick.
- Napp och/eller 30 procents glukos oralt är effektiv smärtlindring för spädbarn.
- EMLA i enstaka doser till nyfödda medför inte biverkningar och är effektivt, dock inte vid hälstick.

Författare

INGE AXELSSON

docent, institutionen för hälsovetenskap, Mitthögskolan, Östersund, och barnkliniken, Östersunds sjukhus.

E-post: inge.axelsson@hvs.mh.se

FAKTARUTA

Adresser till gratis Internettjänster om smärtlindring

1. Databaser nästan helt grundade på randomiserade studier och systematiska översikter

The Cochrane Library nås via Läkarförbundets hemsida: www.slf.se

»The Cochrane Collaboration is an enterprise that rivals the Human Genome Project in its potential implications for modern medicine» [6].

The Oxford Pain Site: www.jr2.ox.ac.uk/bandolier/

En omfattande hemsida som innehåller systematiska översikter över analgetika och andra analgetiska interventioner och en »ligatabell» som rangordnar analgetika efter deras analgetiska förmåga.

Bandolier är en elektronisk gratistidskrift om EBM (evidensbaserad medicin) som ges ut av forskare och läkare i Oxford.

2. Tidskrifter som är pionjärer inom EBM

British Medical Journal (BMJ): <http://www.bmj.com>

Journal of the American Medical Association (JAMA): <http://jama.ama-assn.org/>

Alla artiklar i Läkartidningens kusiner BMJ och JAMA finns i sin helhet gratis på Internet! Referens 2 kan man alltså lätt finna och skriva ut.

Pediatric Pain Letter: <http://is.dal.ca/~pedpain/pplet/pplet.toc.html>

En tidskrift som består av strukturerade, kommenterade sammanfattningar av artiklar av god kvalitet om smärta hos barn. Utges av Dalhousie University, Kanada. Är inte gratis.

3. Databaser som inte är baserade på systematiska litteratursammanställningar

pain.com. A world of information on pain: <http://www.pain.com/index.cfm>

En mycket innehållsrik hemsida.

OMNI (Organising Medical Networked information): <http://www.omni.ac.uk/>

Också en mycket innehållsrik, brittisk hemsida med fackgranskade (peer reviewed) texter. Söker man »pain» får man flera intressanta träffar.

PubMed Central: <http://www.pubmedcentral.nih.gov/>

National Institutes of Health (NIH, USA) har nyligen (april 2000) öppnat denna web-site med avsikten att alla betydande tidskrifter inom »life science» (inklusive biomedicin) ska ges ut i komplett gratisupplaga på PubMed Central! En del stora förlag vägrar medan andra, som BMJ och Critical Care, har lovat att ställa upp. »Sajten» innehåller i september 2000 bara sex tidskrifter, bl a prestigetidskriften PNAS.

Databaserna nåddes på ovanstående hemsidor i september 2000.

die av venpunktion hos nyfödda lindrar 30 procent glukos och 30 procent sackaros oralt smärtreaktionerna lika effektivt, men napp är ännu effektivare [2, 3].

Hälstick

Det finns inga goda skäl att använda hälstick på friska, fullgångna, nyfödda barn. Om hälstick används bör det göras med en automatisk lansett och smärtlindring ges med 1 ml 30-procentig glukos eller sackaros per os [1, 4]. EMLA, paracetamol och bröstmjölk tycks inte minska smärtreaktionerna vid hälstick (data om bröstmjölk är emellertid relativt begränsade). Att värma upp hälen minskar inte smärtreaktionen och gör det inte lättare att få tillräckligt med blod [1].

KOMMENTAR

Översikten [1] är kompletterad med data från två randomiserade studier [2, 4] som inte var publicerade när Ohlsson

och medarbetare gjorde sin sammanställning. Shah och Ohlsson har kommenterat [3] den franska undersökningen [2] som de finner vara av hög kvalitet och som bekräftar tidigare undersökningar av den analgetiska effekten av sockerlösningar och icke-nutritivt sugande (napp).

Prilokainet i EMLA förorsakar bildning av methemoglobin. Nyfödda har låg aktivitet av det enzym som reducerar methemoglobin till hemoglobin. Behandling med EMLA till nyfödda bör därför inte upprepas inom det närmaste dygnet. Hälstick kan bli aktuellt där det inte finns någon som är van vid venpunktion på nyfödda, t ex vid PKU-prov hos distriktssköterskan efter tidig hemgång från BB. Den största smärtan vid hälstick tycks komma från att man klämmer på foten.

I faktarutan finns adresser till Internet-baserade databaser om smärta och behandling av smärta. En konventionell, icke-elektronisk bok med ett icke evidensbaserat brittiskt konsensus om

hur man hanterar smärta hos barn förtjänar också omnämnande [5].

Efter alla år med obedövade hälstick tycks övergången till smärtlindrade venpunktioner ha gått tämligen snabbt och problemfritt på svenska BB och neonatalavdelningar, påskyndad av inhemsk forskning av läkare och sjuksköterskor på Astrid Lindgrens Barnsjukhus och Regionsjukhuset i Örebro.

Referenser

1. Ohlsson A, Taddio A, Jadad AR, Stevens BJ. Evidence-based decision making, systematic reviews and the Cochrane collaboration: implication for neonatal analgesia. In: Anand KJS, McGrath PJ, Stevens BJ. Pain in Neonates. Amsterdam: Elsevier, 2000.
2. Carbajal R, Chauvet X, Couderc S, Olivier-Martin M. Randomised trial of analgesic effects of sucrose, glucose, and pacifiers in term neonates. *BMJ* 1999; 319: 1393-7.
3. Shah V, Ohlsson A. Abstracts from the literature: randomized trial of analgesic effects of sucrose, glucose, and pacifiers in term neonates. *J Pediatrics* 2000; 136: 701
4. Eriksson M, Gradin M, Schollin J. Oral glucose and venepuncture reduce blood sampling pain in newborns. *Early Human Development* 1999; 55: 211-8.
5. Southall D, ed. Prevention and control of pain in children: a manual for health professionals. Royal College of Paediatrics and Child Health. London: BMJ Publishing Group, 1997
6. Naylor CD. Grey zones of clinical practice: some limits to evidence-based medicine. *Lancet* 1995; 345: 840-2.

Cochrane-rapport om karotiskirurgi:

Tidig endarterektomi vid karotisstenos rekommenderas

Sammanfattat

- Karotiskirurgi är granskad med Cochrane-metodik.
- Karotiskirurgi i form av endarterektomi är av nytta för patienter med klar symptomatologi och tät karotisstenos – om sammanlagd morbiditet/mortalitet understiger 6 procent.

I början av 1940-talet disputerade Gösta Hultqvist, sedermera patologiprofessor i Uppsala, på en avhandling med titeln »Über Thrombose und Embolie der Arteria Carotis und hierbei vorkommende Hirnveränderungen».

Denna avhandling hade säkerligen ingen betydelse för karotiskirurgins utveckling, men det är intressant att notera iakttagelsen att arteriosklerotiska lesioner i karotisbifurkationen hade ett samband med cerebrala malacier. Detta utgör i själva verket den patofysiologiska grunden för karotiskirurgi, dvs det kirurgiska avlägsnandet av embolikällan i karotisbifurkationen.

Liten historik

Operationsmetoden började utvecklas i början av 1950-talet på olika ställen; den första publikationen kom från London 1954 [1]. Operationen blev etablerad vid flera centra, medan den hade svårt att finna acceptans vid andra, framför allt på grund av en hög komplikationsfrekvens på en del håll. Försök till randomiserade studier gjordes, men dessa var för små för att ge ett statistiskt underlag [2] eller avbröts i förtid på grund av alltför hög frekvens av postoperativa slaganfall [3].

Trots detta ökade operationsfrekvensen i USA fram till dess att en randomiserad studie visade bristande effekt av transkranieell bypass mellan a carotis externa och interna [4], trots att

studien i fråga egentligen inte hade avsett att analysera karotiskirurgis roll.

Scenen förändrades under 1990-talet, då stora randomiserade multicenterstudier publicerades som gjorde att karotiskirurgi får anses vara en av de bäst utvärderade operationsmetoderna överhuvudtaget. Detta har haft stor betydelse för att operationen nått acceptans. Det rör sig om profylaktisk kirurgi där en fruktad komplikation är postoperativt slaganfall, vilket ju är det tillstånd som man avser förebygga med det kirurgiska ingreppet.

Detta faktum har för övrigt didaktiska implikationer i patientinformations-situationen.

Målsättningen med denna översikt är att kommentera en systematisk litteraturanalys från Cochrane Collaboration avseende karotiskirurgi vid symptomgivande karotisstenos [5]. Analysen finns i huvudsak publicerad även i Journal of Vascular Surgery [6].

Metaanalysens material och metoder

Metaanalysen utgår från en systematisk översikt av randomiserade studier,

Författare

DAVID BERGQVIST
professor i kärnkirurgi,
institutionen för kirurgiska
vetenskaper, enheten för kirurgi,
Uppsala universitet.

där kirurgi jämförs med icke-kirurgi avseende patienter med symptomgivande karotisstenos. Sökstrategin omfattade ett speciellt register över kontrollerade studier inom slaganfallsområdet, Medline från 1966, EMBASE från 1999, Healthstars och Serlines databaser från 1966, Cochrane Controlled Trials Register samt referenslistor i relevanta artiklar.

Effekten mättes som sammanlagt utfall av slaganfall med kvarvarande funktionellt bortfall och/eller död, dvs för patienten kliniskt viktiga resultat. Studiernas validitet avgjordes genom predefinierade kriterier. Metodologisk kvalitet avgjordes avseende såväl design som rapportering. Data extraherades av två personer; vid olika bedömningar beslutades genom konsensusdis-

Tabell I. Studiekarakteristika. Veterans Administration (VA)-studien är utesluten från själva metaanalysen då resultatet också inkluderade crescendo-TIA. Studien hade få relevanta händelser och avbröts när resultat från NASCET offentliggjordes.

	ECST	NASCET	VA
Duration	1981–95	1987–97	1988–91
Antal centra	97	106	16
Ålder, år	63	66	65
(Procent män)	72	69	100
Mediantid mellan sista symtom och randomisering, dagar	52		
Mediantid mellan randomisering och operation, dagar	12	2	
Medeluppföljningstid, år	6,1	5	1

Tabell II. Resultat av metaanalys av ECST och NASCET. Procent och 95 procents konfidensintervall.

Stenosgrad, procent	Relativ riskreduktion	Absolut riskreduktion	Antal att operera för att förebygga ett slaganfall/död
70–99	48 (27–63)	6,7 (3,2–10)	15 (10–31)
50–69	27 (5–44)	4,7 (0,8–8,7)	21 (11–125)
<50	–20 (0–44)	2,2 (0–4,4)	45 (22–∞)

kussion. Data överfördes till ett standardiserat formulär.

Metaanalysens resultat

23 relevanta publikationer från tre randomiserade studier identifierades, samtliga i EMBASE och/eller Medline. De två tidigare nämnda studierna [2, 3] uppfyllde inte inklusionskriterierna. Bedömningen av den metodologiska kvaliteten var god, med ett kappavärde på 0,75 för de båda observatörerna. Data-extraktionsfel låg mellan 1 och 2 procent.

I de tre multicenterstudierna (ECST, NASCET, VA) hade 6 143 patienter randomiserats. Kvalificerande händelse för randomisering var icke-permanent slaganfall och transitoriska ischemiska attacker (TIA). Samtliga studier inkluderade äldre patienter och redovisade distinkta exklusionskriterier. ECST-studien utnyttade dessutom den sk osäkerhets- eller gråzonsprincipen, dvs patienter randomiserades inte när patientens läkare hade en distinkt åsikt angående operationens nytta eller icke-nytta i det enskilda fallet. Gemensamma exklusionskriterier var allvarlig koexisterande sjukdom, möjlig kardiell embolikälla, stort slaganfall, angiografiska tecken på en andra stenosis distalt (tandemlesion) och tidigare ipsilateral karotiskirurgi.

Analyserna var utförda enligt »intention to treat»-principen. Inte någon studie var av förklarliga skäl blindad. Studierna hade något olika definitioner av slaganfall och större slaganfall, »major stroke». I Tabell I sammanfattas basala data om studierna, i Tabell II sammanfattas de viktigaste resultaten. Kirurgins nytta var klart relaterad till stenosgraden. Den relativa riskreduktionen var påtagligt likartad mellan de olika studierna. Av Tabell II framgår också absolut riskreduktion, och det som kanske för en kliniker är mest intressant, nämligen hur många patienter som behöver opereras för att förebygga utvecklingen av slaganfall och/eller död hos en patient. Vare sig man analyserar död eller stort slaganfall, dödligt eller invalidiserande slag eller ipsilateralt slaganfall utfaller resultatet till kirurgins förmån på ett statistiskt signifikant sätt.

Mätning av livskvaliteten

Från NASCET-studien finns också en mer detaljerad analys av det funktionella utfallet av en livskvalitetsmätning [7]. Man använde ett tidigare validerat



Den första cerebrala angiografien genomfördes av den portugisiska neurologen Egas Moniz 1927. Han fick senare Nobelpriset för lobotomi som behandlingsmetod. Karotisangiografi ligger till grund för stenosgradering i ECST och NASCET, de stora randomiserade studier som visat värdet av kirurgi vid karotisstenos.

instrument som mätte några olika domäner, bl a syn, språkförståelse, talförmåga, sväljningsförmåga, funktion i extremiteter samt integrerade funktioner såsom förmåga att gå ut och handla eller att gå ut överhuvudtaget. Med tiden hade den opererade gruppen statistiskt signifikant bättre resultat. Elva patienter måste opereras för att förhindra att en tidigare icke-handikappad person skulle drabbas av slaganfall, med gravt handikapp som följd.

Allvarliga komplikationer var ovanliga

Allvarliga komplikationer, förutom slaganfall, var ovanliga: hjärtinfarkt (1–2 procent), sårinfektion (3 procent), hematom (5 procent), reversibel perifer nervskada (5–7 procent). Under de första 30 dagarna var risken för slaganfall störst i operationsgruppen. I en subgruppsanalys från ECST-studien med multivariat angreppssätt påvisades ökande effekt vid tilltagande stenosgrad, stigande ålder och manligt kön. Dessa faktorer var gynnsamma också i NASCETs originalpublikation från 1991.

Metaanalysens diskussion och konklusioner

I metaanalysen har man slagit samman data från studier av liknande men inte identisk design och inte riktigt identiska effektmått. Man framhåller att studierna är stora och väl genomförda samt att de analyserar för patient och samhälle viktiga resultat. Man påpekar att operationsgruppen hade låga komplikationsfrekvenser avseende slaganfall och död, och poängterar det viktiga i att centra som opererar karotisstenos på samma indikationer som i studierna måste monitorera sina resultat prospektivt för att visa att man håller sig på en liknande nivå som i de randomiserade studierna.

Man påpekar att randomiseringen gjordes först efter angiografisk dokumentation av stenosgrad, och att man måste ta hänsyn till den komplikationsrisk som belastar angiografimetoden. Sedan 1990 finns fyra prospektiva studier som analyserar komplikationsrisken efter angiografi hos patienter med symtomatisk cerebrovaskulär sjukdom. Den sammanlagda risken för permanent slaganfall eller död var i dessa studier 1,1 procent (95 procents konfidensintervall 0,6–2,0).

En rimlig slutsats blir därför att patienter inte skall angiograferas förrän man verkligen bestämt sig för att indi-

Kortikosteroider effektiva vid behandling av pseudokrupp

Metaanalys bekräftar svensk behandlingstradition

Nyligen publicerades i BMJ [1] en metaanalys av effekten av kortison vid pseudokruppbehandling. Man rapporterar genomgående positiva erfarenheter, vilket givetvis är kliniskt betydelsefullt att få bekräftat i en undersökning av denna dignitet.

Artikeln är dessutom av extra stort intresse för svenska läkare, eftersom de första behandlingsförsöken av pseudokrupp med kortison utfördes i Sverige.

Krupp vanligt hos barn

Andningssvårigheter i samband med luftvägsinfektioner är förhållandevis vanligt förekommande hos barn upp till 6 års ålder. Orsaken är svullnad av slemhinnorna omedelbart under stämbanden, subglottisk laryngit, som mer eller mindre blockerar luftpassagen genom trakea och ger en inspiratorisk stridor och skrällande hosta. I svårare fall uppkommer indragningar i torax i samband med inandning, och – när andningssvårigheterna är särskilt uttalade – cyanos och risk för kolsyrenarkos.

Tillståndet benämns i anglosaxisk litteratur »croup», i svensk litteratur pseudokrupp, att skilja från »äkta» krupp, dvs andningssvårigheter vid difteri.

Liten historik

Den allmänt accepterade terapin under 1950- och 1960-talen var luftfukt-

ning med tält över sängen och luftfuktare, hostdämpande medicin med efedrin samt antibiotika. Patienterna övervakades noga. Vid påtaglig svikt i andning och syresättning trakeotomerades patienterna.

Vid 1950-talets början prövade Mårtensson och medarbetare i Borås att använda prednisolon som terapi vid pseudokrupp. De noterade förbättringar av patienternas tillstånd. I en stor placebo-kontrollerad studie 1960 visades förbättring av symtomen inom en timme efter insatt behandling. Det förelåg en statistisk skillnad mellan behandlade fall och placebofall med avseende på symtomens duration [2]. Författarna påpekade att resultaten var mer påtagliga hos patienter med allergiska symtom i anamnesen.

Resultaten efter behandling av pseudokrupp med kortikosteroider var så uppmuntrande att Mårtensson och medarbetare rekommenderade att kortikosteroider skulle ges till alla pseudokrupspatienter med symtom på obstruktion av andningsvägarna.

Kortikosteroidbehandling vid pseudokrupp togs därför redan på 1960-talet upp som en standardbehandling vid många öron-, näs- och Halskliniker i Sverige. Antibiotikabehandling vid subglottisk laryngit byggde på antagandet att bakteriell infektion hade del i utvecklingen av symtomen. Virologiska och bakteriologiska undersökningar under 1960-talet visade emellertid att flertalet patienter hade en virusinfektion, vanligen parainfluenza typ 1, 2 och 3, medan patologiska bakterier inte påvisades oftare än hos en kontrollgrupp [3]. Indikationen för behandling med antibiotika var därför tveksam. Senare har även visats att luftbefuktning inte har någon betydelse vid behandling av pseudokrupp [4].

Nya studier på 1980- och 1990-talen

Under 1980- och 1990-talen har en rad kontrollerade studier utförts av ef-

Uppläggning av den kommenterade artikeln

(The effectiveness of glucocorticosteroids in treating croup: meta-analysis. BMJ 1999; 319: 595-600):

Artikelns frågeställning: Är kortikosteroider effektiva vid behandling av pseudokrupp?

Design: Metaanalys av randomiserade kontrollerade studier.

Använda parametrar: Symtomskala mätande kruppanfallets svårighetsgrad, tilläggsbehandling (adrenalin, epinefrin, antibiotika eller ytterligare kortikosteroider), sjukhusvistelsens längd och antal inläggningar.

Resultat: 24 studier uppfyllde kriterierna och inkluderades. Glukokortikoidbehandling gav symptomförbättring efter sex och tolv timmar. Efter 24 timmar var förbättringen inte längre signifikant. Barnen behövde också mindre tilläggsbehandling med adrenalin och deras vistelse på sjukhus var signifikant kortare.

Slutsatser: Dexametason och budesonid minskar kruppsymtomen redan sex timmar efter insatt behandling. Mindre tilläggsbehandling behövs och sjukhusvistelsen förkortas hos kruppbarn som fått kortison.

fekten av kortikosteroider vid pseudokrupp. Flertalet av dem visar en statistiskt signifikant effekt på både symtom och symtomduration [5, 6].

Ett antal av dessa kontrollerade studier sammanfördes av Monica Aulsebrook och medarbetare [1] till en metaanalys av glukokortikoidbehandling och dess effektivitet vid behandling av pseudo-

Författare

JAN WERSÄLL
professor emeritus

LOTTA HESSÉN SÖDERMAN
biträdande överläkare

RICHARD KUYLENSTIERNA
docent, överläkare; samtliga ÖNH-kliniken, Karolinska sjukhuset, Stockholm.

krupp. Kraven för selektion till meta-studien var dels en gradering av symtomen och eventuella symptomförbättringar, dels att en behandling med någon form av glukokortikosteroid skulle vara jämförd med placebo eller annan form av behandling i en randomiserad dubbelblind studie.

Mer än hälften av de kontrollerade studier som sammanställdes i metaanalysen användes för klassifikation av symtomen på Westleys 17-gradiga skala som anger luftpassage (2 poäng), stridor (2 poäng), interkostala indragningar (3 poäng), cyanos (5 poäng) och medvetandegrad (5 poäng). Av de 24 arbeten som studerades i metaanalysen hade 17 utvärderat dexametason, nio budesonid och tre metylprednisolon.

Resultat av metastudien

Resultatet av metastudien visade en förbättring av symtomen sex timmar och tolv timmar efter insatt terapi, medan förbättringen efter 24 timmar inte längre var signifikant. Behovet av adrenalinbehandling minskade efter glukokortikosteroidbehandling. Antalet timmar som barnen tillbringade i intensivbehandling minskade också, liksom antalet timmar på sjukhus för de inneliggande barnen.

Konklusionen blev att alla barn med andningssvårigheter vid pseudokrupp bör behandlas med kortikosteroider i någon form.

Adrenalin som alternativ

Alternativet till glukokortikoidbehandling har varit adrenalin, främst som inhalationsbehandling i aerosol av L-adrenalin, eller racemiskt adrenalin antingen ensamt eller i kombination med kortikosteroider [7]. Inhalerat adrenalin ger en minskning av svullnaden i trakeaslemhinnan, ofta en minskning av andningsmotståndet och en förbättring av symtomen. Effekten är emellertid kortvarig och försvinner efter cirka två timmar, varefter det finns risk för reci-

div av symtomen. I en randomiserad dubbelblind studie, där dexametason jämförts med adrenalininhalation respektive placebo, har dexametason och adrenalinaerosol bättre effekt än placebo på symtomen vid pseudokrupp, men adrenalineffekten är mer kortvarig än dexametason-effekten [8, 9]. Upprepade adrenalinbehandlingar ökar risken för komplikationer i form av takykardi och arytmier.

Det har rekommenderats att barn som behandlats med inhalation av racemiskt adrenalin skall kvarstanna på sjukhus med resurser att behandla eventuella komplikationer under åtminstone sex timmar efter behandlingen [10].

Slutsats och rekommendation

Den genomgång av ett stort antal studier avseende resultatet av kortisonbehandling av barn med pseudokrupp som Ausejo och medarbetare utfört i sin metaanalys ger vid handen att alla barn med måttliga eller svåra andningssvårigheter på grund av pseudokrupp i första hand bör behandlas med kortikosteroidpreparat.

Sådan behandling, utan vare sig luftbefuktning eller antibiotika vilkas effekt ej kunnat dokumenteras i detta sammanhang, synes vara den adekvata förstahandsterapi vid pseudokrupp. Inhalationsbehandling med adrenalin har också en plats i terapiarsenalen, men dess effekt är inte lika långverkande som kortisonets.

Vår rekommendation blir således att fortsätta med den i Sverige vanligaste behandlingen, dvs betametason i form av vattenlösliga tabletter (Betapred 0,5 mg): 8 tabletter för barn mindre än 10 kg, 12 tabletter för barn som väger mer än 10 kg. Tabletterna ges som engångsdos. Om barnet krävs rekommenderas i enlighet med metaanalysens resultat [11-13] budesonid i inhalationsform (Pulmicort) eller kortison intramuskulärt, t ex Decadron 0,3mg/kg eller Solu-Cortef 25-50 mg.

Referenser

1. Ausejo M, Saenx A, Pham B, Kellner JD, Johnson DW, Moher D et al. The effectiveness of glucocorticosteroids in treating croup: meta-analysis. *BMJ* 1999; 319: 595-600.
2. Mårtensson B, Nilsson G, Torbjär JE. The effect of corticosteroids in the treatment of pseudo-croup. *Acta Otolaryngol (Stockh) Suppl* 1960; 158: 62-71.
3. Svartz-Malmberg G, Andrén E, Carlström G, Wadström T, Wersäll J. Synpunkter på virusdiagnostik vid subglottisk laryngit (pseudokrupp). *Läkartidningen* 1968; 65: 141-8.
4. Bourchier D, Dawson KP, Fergusson DM. Humidification in viral croup, a controlled trial. *Australian Pediatric Journal* 1984; 20: 289-91.
5. Geelhoed GC. Croup; new perspectives on corticosteroid therapy. *Clin Immunother* 1996; 5: 260-7.
6. Kairys SW, Olmstead EM, O'Connor GT. Steroid treatment of laryngotracheitis; a meta-analysis of the evidence from randomized trials. *Pediatrics* 1989; 83: 683-93.
7. Garner HG, Powell KR, Roden VJ, Cherry JD. The evaluation of racemic epinephrine in the treatment of infectious croup. *Pediatrics* 1973; 52: 68-71.
8. Westley CR, Cotton EK, Brooks JG. Nebulized racemic epinephrine by IPPB for treatment of croup; a double-blind study. *Am J Dis Child* 1978; 132: 484-7.
9. Waissman Y, Klein BL, Boenning DA, Young GM, Chamberlain JM, O'Donnell R et al. Prospective randomized double-blind study comparing L-ephedrine and racemic epinephrine aerosol in the treatment of laryngotracheitis (croup). *Pediatrics* 1992; 89: 302-5.
10. Dawson K, Kennedy D, Asher I, Cooper D, Cooper P, Francis P. The management of acute laryngo-tracheobronchitis (croup): A consensus view. *J Paediatr Child Health* 1992; 28: 223-4.
11. Klassen TP, Feldman ME, Watters LK, Sutcliffe TRN, Rowe PC. Nebulized budesonide for children with mild-to-moderate croup. *N Engl J Med* 1994; 331(5): 285-9.
12. Hessén-Söderman L, Kuylenstierna R. Inhalationssteroider vid pseudokrupp. Goda resultat i Kanada, svensk tillämpning dröjer. *Läkartidningen* 1995; 92: 268.
13. Linder A. Från empiri till vetenskaplig säkerhet vid falsk krupp. God dokumentation för kortikosteroider. *Läkartidningen* 1994; 91: 3169-72.

Dyspepsi samlingsbegrepp för en av de vanligaste folksjukdomarna

Statens beredning för medicinsk utvärdering (SBU) har utkommit med en rapport, »Ont i magen», i vilken dagens kunskap om metoder för diagnos och behandling av dyspepsi har samlats. Rapportens sammanfattning presenteras här i något förkortad form.

Magbesvär utgör efter huvudvärk och trötthet en av de vanligaste anledningarna till läkarbesök; dyspepsi står för omkring 3 procent av dessa besök. Dyspepsi är ett sammanfattande begrepp för ihållande (veckor till månader) smärta eller obehag lokaliserade till övre delen av buken. Ordet dyspepsi kommer från grekiskan och betyder dålig (dys) matsmältning (pepsi).

Bland flera orsaker till dyspepsi märks ulkussjukdom, dvs sår lokaliserade till magsäck eller tolvfingertarm. Upptäckten av att ulkussjukdom är relaterad till en infektion av bakterien *Helicobacter pylori* har i förening med utvecklingen av effektiva läkemedel mot magsårssjukdom markant förändrat behandlingsstrategin, med förbättrade resultat som följd.

Detta har fött förhoppningen att övriga orsaker till dyspepsi skulle kunna behandlas framgångsrikt på likartat sätt. Kontroversiella uppfattningar härvidlag, liksom markant ökad förskrivning av potenta syrahämmande medel, utgjorde några av incitamenten till SBU-rapporten.

Vad orsakar dyspepsi ?

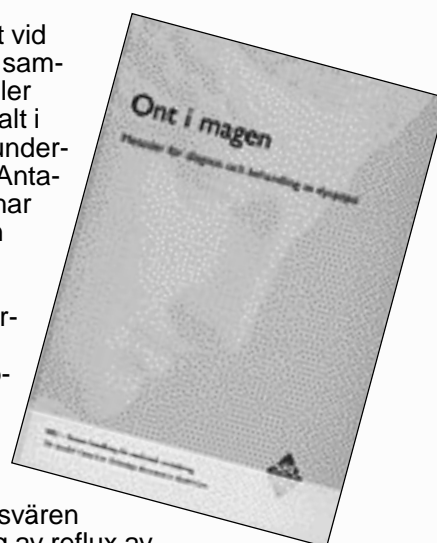
Dyspepsi har många orsaker och är antingen organiskt eller funktionellt betingad. Organisk dyspepsi innefattar framför allt ulkussjukdom, dvs sår i tolvfingertarm eller magsäck, inflammation i matstrupe (esofagit) samt cancer i matstrupe och magsäck. I rapporten tas endast upp samband mellan *H pylori* och magsäckscancer. När organiska förändringar inte kan påvisas föreligger refluxsjukdom eller funktionell dyspepsi.

Ulkussjukdom

Uppkomst av sår i magsäcken och tolvfingertarmen har sedan lång tid uppfattats sammanhånga med en störd balans mellan aggressiva (enzymet pepsin och saltsyra) och skyddande faktorer. Behandling av ulkussjukdom har därför huvudsakligen varit inriktad på

Sammanfattat

- Största informationsvärdet vid utredning av dyspepsi, ett samlingsbegrepp för smärta eller obehag lokaliserade centralt i övre delen av buken, har undersökning med gastroskop. Antalet årliga undersökningar har ökat markant. Så har även läkemedelskostnaderna.
- Mer än hälften av patienterna med dyspepsi får diagnosen funktionell dyspepsi, vilket innebär att någon orsak till besvären inte kunnat fastställas. Symtomlindring kan uppnås med syrahämmare när besvären har ett dominerande inslag av reflux av syra till matstrupen.
- Majoriteten av patienter med ulkussjukdom (sår i magsäck och tolvfingertarm) är kroniska bärare av bakterien *Helicobacter pylori*, som orsakar inflammatoriska förändringar i magsäcken. En mycket stor andel av befolkningen bär på denna infektion, men endast en av sex drabbas av ulkus. Det är därför inte meningsfullt att införa screening för bärare av *H pylori*-infektion
- Endast patienter med ulkussjukdom skall behandlas med antibiotika riktade mot bakterien *H pylori*. Behandling med läkemedel som hämmar bildningen av saltsyra i magsäcken, s k protonpumpshämmare, i kombination med två antibiotika både läker såren i tolvfingertarmen och minskar risken för återfall. Patienter med refluxsjukdom blir snabbt symtomfria med protonpumpshämmare.
- NSAID-preparat medför ofta biverkningar i form av dyspepsi och utgör en vanlig orsak till ulkus, framför allt i magsäcken.
- Svåra och alarmerande magbesvär, exempelvis blödning, bör föranleda akut eller snar undersökning av matstrupe och magsäck med gastroskop.
- Magsäckscancer, som svarar för mindre än 1 procent av all dyspepsi, förekommer sällan hos individer under 45 års ålder. Med *H pylori*-infektion föreligger något förhöjd risk att drabbas av magsäckscancer.



att motverka bildning av saltsyra eller att neutralisera denna. En vanlig uppfattning är att olämplig kost, stort kaffeintag och psykologisk stress har betydelse för uppkomst av ulkussjukdom. Det vetenskapliga stödet är emellertid begränsat, vilket beror på metodologiska problem. Rökning har visats bidra till uppkomst av ulkus och försämra läkning av sår. För alkohol är uppgifterna motsägelsefulla; såväl förbättrad som försämrad sårhäkning har rapporterats.

För ett femtontal år sedan gjordes den betydelsefulla upptäckten att majoriteten av patienter med ulkussjukdom uppvisade inflammatoriska förändringar (gastrit) i magsäcken som var orsakade av *H pylori*. Denna bakterieinfektion är spridd över hela världen och förekommer i väsentligt högre grad i västvärlden än i utvecklingsländerna, vilket pekar på att socioekonomiska förhållanden har stor betydelse för smittspridning. Smittvägarna är alltså inte kända. *H pylori* förekommer oftare i högre ålder; blodprov pekar på att 40–50 procent av 50-åringarna i Sverige är bärare.

Refluxsjukdom

Typiska besvär vid refluxsjukdom är halsbränna, »bröstbränna» (smärta bakom bröstbenet) och sura uppstötningar, som uppkommer vid reflux av saltsyra och annat maginnehåll till matstrupen. Hos inemot hälften av patienterna blir delar av matstrupen inflammatoriskt förändrade eller säte för ytliga sårbildning av olika svårighetsgrad.

Funktionell dyspepsi en uteslutningsdiagnos

Funktionell dyspepsi är en uteslutningsdiagnos, dvs när noggrann utredning av en patient med dyspepsi inte påvisat någon organisk orsak eller någon rubbning av mag-tarmkanalens funktion, exempelvis reflux. Karaktären av besvär, liksom den subjektiva upplevelsen, uppvisar betydande individuell variation.

Ulkus framkallat av NSAID-läkemedel

Läkemedel som ingår i NSAID-gruppen (anti-inflammatoriska medel av icke-kortison typ) medför ofta dyspepsi och ibland sår som oftast är lokaliserade till magsäcken. Preparaten används såväl vid kroniska led-muskelsjukdomar, t ex ledgångsreumatism, som vid akuta smärttillstånd i leder och muskler. Symtomen korrelerar dåligt med förändringarna i magsäckens slemhinna, vilket sammanhänger med NSAID-medlens smärtstillande effekt. Blödning, i enstaka fall livshotande, och anemi är inte sällan de första klinis-

ka tecknen på ulkus. Biverkningarna sammanhänger med att läkemedlen försvagar skyddsmekanismerna för magsäckens slemhinna och gör denna mer sårbar för kroppsegna substanser, t ex saltsyra, och yttre faktorer. Hög dosering av NSAID-preparat, hög patientålder, samtidig behandling med kortison och tidigare ulkussjukdom utgör riskfaktorer.

Alarmerande fynd av ett stort antal fall med dödlig utgång i USA, där överanvändningen av NSAID förefaller vara avsevärd, motiverar ökad information om risken för biverkningar.

ASA (acetylsalicylsyra) som k lågdos har fått stor användning som sekundär och tertiär profylax vid hjärt-kärlsjukdom (kärlkramp, hjärtinfarkt och slag-anfall). Risken för magbesvär är visserligen liten, men allvarliga magblödningar kan förekomma, även hos personer som tidigare inte haft ulkussjukdom.

Helt nyligen har i Sverige registrerats NSAID-preparat som är anti-inflammatoriskt väsentligt mer verksamma än vad de nedsätter slemhinnans skyddsmekanismer. Dessa medel, s k Cox-2, är teoretiskt tilltalande. Den vetenskapliga dokumentationen är idag fortfarande begränsad. Preparatens ställning i den terapeutiska arsenalen, i första hand vid kroniska reumatiska sjukdomar, kan antas bli avhängig av kostnad i relation till såväl långsiktiga effekter som eventuella biverkningar.

Förekomst av dyspepsi

Epidemiologiska studier av sjukdomar som ingår i begreppet dyspepsi är förknippade med metodologiska problem, bl a avsaknad av objektiva undersökningsfynd hos många drabbade och resursproblem att med noggrann metodik undersöka stora, representativa populationer.

Vad som genomgående framkommer i studierna är att tillståndet är vanligt förekommande. Under en tremånadersperiod har t ex i storleksordningen 30 procent av den vuxna befolkningen haft dyspepsi. Endast en mindre andel av dem med besvär har ulkus (1 av 10) eller refluxsjukdom (1 av 7), vilket innebär att funktionell dyspepsi är den dominerande inom diagnosgruppen.

Vissa könsskillnader har konstaterats. Ulkus i tolvfingertarmen förekommer oftare hos män, medan funktionell dyspepsi oftare förekommer hos kvinnor. I en begränsad populationsbaserad studie från Sverige uppvisade yngre kvinnor den högsta frekvensen av läkarbesök på grund av dyspepsi. Magsäckscancer, som sällan förekommer under

40 års ålder, är ungefär dubbelt så vanlig hos män som hos kvinnor.

Symtomatologi vid utredning av dyspepsi

Det är i regel inte möjligt att enbart på bas av magbesvärens karaktär med rimlig säkerhet fastställa orsaken till dyspepsi hos en viss patient. Däremot kan anamnesen göra olika diagnoser mer eller mindre sannolika och underlätta valet av bekräftande eller uteslutande undersökningar. Halsbränna och sura uppstötningar talar således starkt för refluxsjukdom. Symtomens duration är också en viktig faktor i utredning och behandling av patienter med dyspepsi, liksom besvärens förlopp.

De symtom som talar för respektive emot olika diagnoser ingår i den kunskap som läkare på alla nivåer förutsätts ha. Tillämpningen av denna kunskap varierar dock. Det finns därför behov av klarare definitioner av symtom eller symtomkonstellationer som föranleder ett visst diagnostiskt eller terapeutiskt handlande. Vid dyspepsi brukar man tala om s k alarmerande symtom (t ex viktminskning, svart avföring, blodig kräkning, sväljningssvårigheter), som alltid skall föranleda prioriterade undersökningar.

Ytterligare ett sådant alarmerande symtom är nydebuterad dyspepsi hos patienter som är äldre än 45 år.

Gastroskopi, pH-mätning och ultraljud

Av de undersökningar som kan företas i syfte att klargöra den underliggande diagnosen vid dyspepsi intar gastroskopin en särställning.

Mellan 10 och 20 procent av patienterna med mer än tillfällig dyspepsi har ulkus, ca 10 procent har refluxesofagit, medan <1 procent har cancer. Dessa tre diagnoser ställs med hög säkerhet med hjälp av gastroskopi.

24-timmars pH-mätning i esofagus har sannolikt det näst högsta informationsvärdet i utredningen av patienter med dyspepsi och antas ge underlag för diagnosen gastroesofageal refluxsjukdom hos upp till 70 procent av patienterna med ihållande dyspepsi.

Ultraljud är av värde för att diagnostisera sjukliga processer i gallvägar, bukspottkörtel och buk, vilka förekommer hos cirka 5 procent av patienterna med dyspepsi.

H pylori-diagnostik

Rutinmetoderna i primärvården är som regel icke-invasiva (kräver inte gastroskopi) och kan t ex bestå av att antikroppar mot *H pylori* påvisas i helblod (snabbtest = svar direkt) eller i serum som skickas till laboratorium. Ett säkrare prov är ureasutandningstestet,

som bygger på att *H pylori* har egenskapen att med hjälp av enzymet ureas kemiskt spjälka tillfört urea.

Motsvarande undersökning utförs ofta på vävnadsprov, som genom biopsi kan erhållas i samband med gastroskopi (invasiv undersökning). Vid gastroskopi kan vävnadsprov avlägsnas för odling av *H pylori* respektive för mikroskopisk undersökning av magsäcks-slemhinnan samt för ett s k snabb-ureastest.

Projektets metodik: litteraturgranskning

För sammanställning av kunskapsläget avseende behandling av dyspepsirelaterade sjukdomar utgjorde en kritisk granskning av vetenskapliga studier det viktigaste steget. Litteraturgenomgången, baserad på Medline för perioden 1966–1997 inbegrep huvudsakligen randomiserade kontrollerade studier (RCT), som förmår isolera behandlingseffekterna från inflytandet av ovidkommande faktorer. Med tillämpning av en strukturerad mall med olika kvalitetskriterier identifierade arbetsgruppen studier för den fortsatta granskningen och sammanställningen.

Beslutsanalys

I rapporten tillämpas beslutsanalysen i form av trädidiagram. Analysen bygger på statistisk sannolikhetslära (Bayes' teorem angående betingad sannolikhet). Syftet med analysen är att skatta den förväntade kostnaden och de förväntade hälsokonsekvenserna av olika handlingsalternativ. Därmed kan utrednings- och behandlingsalternativen jämföras med avseende på kostnaden per uppnått resultat, exempelvis kostnaden per kvalitetsjusterat levnadsår.

Utredning av gastrokopiverksamheten

En landsomfattande enkätundersökning av gastrokopiverksamheten, bl a antal utförda undersökningar och väntetider, företogs för år 1997.

Resultat av litteraturgranskningen

Ulkussjukdom. De övergripande fynden är att korttidsbehandling (1–2 veckor) med syrahämmare läker minst 90 procent av sårerna i tolvfingertarmen utan att *H pylori*-infektionen påverkas. Om behandlingen kombineras med antibiotika riktade mot *H pylori*-infektionen minskar dessutom risken för återfall avsevärt. De gynnsammaste resultaten har uppnåtts med protonpumpshämmare i kombination av amoxicillin, klaritromycin och metronidazol (s k trippelterapi). De i litteraturen vanligaste trippelbehandlingarna innehåller en kombination av protonpumpshämmare med

amoxicillin och klaritromycin eller klaritromycin och metronidazol.

Dokumentationen är sämre när det gäller sår i magsäcken, avseende såväl läkning som återfall. Genomgående fynd är att enbart syrahämmare medför sårhäkning inom 6–8 veckor i omkring 80 procent av fallen. Återinsjuknandet under det första året kan reduceras markant, från omkring 50 procent till 5 procent, om motsvarande behandling som vid tolvfingertarmsår genomförs. Blödningsskomplikationer till följd av ytliga sårbildningar och ulkus förefaller likaledes att reduceras med den behandlingsstrategi som numera används.

Erfarenhetsmässigt tenderar antalet terapeutiska misslyckanden i klinisk praxis att ligga högre än i kliniska prövningar, som alltid omfattar väldefinierade studiepopulationer. Detta kan främst tillskrivas en mindre god följsamhet, »compliance», till medicineringen. Ett annat skäl kan vara att *H pylori* är resistent mot insatt antibiotikum. Det saknas studier som belyser värdet av olika strategier vid behandling av duodenalsår som inte läkt ut (ca 5–10 procent).

Sammanfattningsvis är trippelterapi den mest effektiva behandlingen av *H pylori*-positiva patienter med sår i tolvfingertarmen eller i magsäcken. Behandlingen består av syrahämmande medel i kombination med två antibiotika. Detta stämmer väl med de riktlinjer som Läkemedelsverket utfärdat. Förskrivningen av de aktuella läkemedlen förefaller vara i överensstämmelse med riktlinjerna, men det är inte närmare klarlagt hur god följsamheten är. (Efter rapportens publicering har nya rön publicerats. Se addendum nedan.)

Refluxsjukdom. Med läkemedel som hämmar saltsyrabildningen kan esofagiten läkas ut och symtomen lindras. Den terapeutiska effekten är relaterad till grad och duration av syrahämning, som bör vara effektivare än vid ulkusterapi.

Eftersom diagnosen esofagit endast kan ställas med gastroskopi, och symto-

met halsbränna är vanligt förekommande, krävs en behandlingsstrategi för patienter när diagnosgrundande undersökning av olika skäl inte bör eller kan ifrågakomma. Litteraturgranskningen har därför speciellt gällt endoskopinegativ gastroesofageal refluxsjukdom. Många studier uppvisar brister med avseende på karakteristik av patienterna, såsom symtomens varaktighet och svårighetsgrad. Placeboeffekten är mycket hög, omkring 40 procent, varför den symtomlindrande effekten med medel som hämmar bildningen av saltsyra, främst av gruppen protonpumpshämmare, inskränker sig till omkring 30 procent.

Sammanfattningsvis ger syrahämmande medel bättre symtomlindring än placebo vid refluxsjukdom även om esofagit inte föreligger.

Funktionell dyspepsi. De kliniska studierna omfattar som regel grupper av patienter med skilda symtomkomplex. I brist på objektiva fynd vid funktionell dyspepsi utvärderas behandlingseffekten ofta med hjälp av skalor för subjektiv skattning av dels magbesvären, dels patientens livskvalitet.

De två principiellt olika behandlingsstrategier som studerats är dels eradikering (borttagande) av *H pylori*, dels symtomlindring med läkemedel riktade mot förmodade rubbningar av magsäckens olika funktioner.

Först under de två senaste åren har studier publicerats som varit kontrollerade (RCT) och i övrigt väl utformade samt omfattat tillfredsställande stora patientmaterial. Endast en av dessa studier styrker att behandling med antibiotika av *H pylori* medför bättre symtomlindring än placebo. För samtliga fyra studier gäller att endast en av fem patienter som erhöll aktiv behandling var symtomfri vid uppföljningstidens slut. Dessa resultat talar starkt emot att eradikering (eliminering) av *H pylori* har någon plats i behandlingen av funktionell dyspepsi.

Denna slutsats får stöd av kunskapen att patienter med funktionell dyspepsi är bärare av *H pylori*-infektion i samma eller obetydligt högre frekvens än personer som inte har motsvarande magbesvär. Det råder dessutom inget samband mellan symtomen och de inflammatoriska förändringar i magsäcken som orsakas av *H pylori*. I avvaktan på väl designade studier och evidensbaserad kunskap om sjukdomsorsak utgör funktionell dyspepsi ingen indikation för behandling med syrahämmande medel.

Den slutsats man kan dra av dessa kliniska studier är att inget medel varit mer effektivt än placebo. Orsaken härtill är främst att vår kunskap om orsaker

Addendum om ulkussjukdom
Duodenal- och ventrikelsår. Behandling med PPI och två antibiotika under 1 vecka eradikerar *H pylori* i ca 90 procent. För duodenalsår räcker en veckas s k trippelbehandling utan efterföljande syrahämmande behandling för att uppnå läkning i >90 procent (nytt efter SBU-rapporten). Efterföljande syrahämmande behandling med PPI i 5–7 veckor rekommenderas i dagsläget för ventrikelsår för att uppnå läkning.

till funktionell dyspepsi, ett heterogent symtomkomplex, är så ofullständig.

Ulkus framkallade av NSAID-läkemedel. En huvudfråga är om ulkus i anslutning till NSAID-användning förstärks vid samtidig H pylori-infektion, med andra ord om eradikering är en framkomlig väg att förhindra uppkomst av ulkus. De tre av totalt fyra studier som är metodologiskt invändningsfria har olika uppläggning och skillnader i patientmaterial. De visar ändå entydigt att förekomst av H pylori-infektion inte inverkar på läkning av ulkus eller återfall av ulkus orsakade av NSAID, oavsett om dessa är lokaliserade till magsäck eller tolvfingertarm. I den fjärde placebokontrollerade studien medförde eradikering av H pylori med trippelterapi, vari ingick vismut, däremot signifikant lägre ulkusuppkomst under pågående NSAID-behandling.

Sammantaget saknas stöd för att eradikeringsbehandla patienter inför eller under NSAID-medicinering. När NSAID-medicinering insätts bör preparat med minsta biverkningsrisk väljas; vidare bör lägsta effektiva dosering användas. Vid vissa kroniska ledsjukdomar utgör smärtstillande läkemedel utanför NSAID-gruppen inte sällan lämpliga alternativ.

Oberoende av val av syrahämmande läkemedel läker sårn bättre om medicinering med NSAID avbryts. Detta bör följaktligen, om möjligt, vara den första åtgärden. I många fall är det nödvändigt med hänsyn till grundsjukdomens natur att fortsätta NSAID-behandlingen. Likaså kan profylaktisk behandling övervägas dels när medicinering med NSAID inleds eller pågår hos patienter som löper ökad risk för ulkus (primärprofylax), dels till patienter som redan varit drabbade av ulkus under behandling med NSAID (sekundärprofylax).

Fyndet i de fåtal studier som accepterats för granskning i rapporten visar att samtliga behandlingsprinciper medför sårsläkning. Protonpumpshämmaren omeprazol har i två jämförande studier visat signifikant bättre effekt än misoprostol respektive histamin-2-receptorantagonisten ranitidin. Hos patienter som löper ökad risk för NSAID-biverkning från magsäcken skall förebyggande behandling övervägas. I de primär- och sekundärprofylaktiska studierna redovisas i princip samma resultat som i läkningsstudierna, men de skiljer sig något åt beroende på sårets lokalisering. Omeprazol har en bättre skyddande effekt än både misoprostol vad avser sår i tolvfingertarmen och ranitidin vad gäller sår i magsäcken. Misoprostol förhindrar uppkomst av sår i magsäcken mer effektivt än ranitidin.

Magsäckscancer och H pylori. Magsäckscancer uppträder oftare hos personer som är bärare av H pylori än hos icke-infekterade. Detta sammanhänger med att de inflammatoriska förändringarna i magsäckens slemhinna kan utgöra ett förstadium till cancer. Riskökningen är dock ringa, vilket betyder att andra för närvarande inte kända faktorer samverkar. Hit hör att vissa H pylori-stammar kan vara mer predisponerade än andra.

I nuläget är det inte möjligt att identifiera vilka bärare av H pylori-infektion som löper risk för magsäckscancer. Följaktligen finns ingen grund för screening av H pylori-infektion i cancerpreventivt syfte.

Utredning av gastrokopiverksamheten

Av en enkätundersökning avseende 1997 framgår att 1 748 gastrokopier per 100 000 invånare utförts, vilket innebär en nära 40-procentig ökning jämfört med för tio år sedan. En mindre andel företas vid akuta tillstånd, såsom blödningar. Anledningen till att betydande geografiska skillnader föreligger är inte känd, dvs i vilken utsträckning som indikationer divergerar eller andra omständigheter existerar är inte klarlagt. Väntetiden för fall som varken är akuta eller i övrigt icke är prioriterade har minskat något. Mediantiden är för närvarande 3,5 veckor, men spännvidden är betydande.

Resultat av beslutsanalys

Med hjälp av beslutsanalys har »test and scope»-strategin (testa först, gastrokopiera dem som inte blir bra efter behandling) jämförts med en strategi som innefattar prompt gastroskopi och riktad behandling. Analysen omfattar personer som för första gången söker för dyspepsi, är 45 år eller yngre, inte har sura uppstötningar eller halsbränna som dominerande besvär och inte har några alarmerande symtom som vikt-nedgång, blödningar från mag-tarmkanalen, gulsot, eller palpabel tumör. I beslutsmodellen följs patienter med avseende på utredning, behandling, behandlingseffekter och kostnader under ett år från första läkarkontakten.

Även om beslutsanalysen endast är en matematisk analys kan vissa slutsatser dras. Den helt avgörande frågan för test and scope-strategins framgång är hur de patienter som med hjälp av H pylori-testet undandras vidare utredning reagerar. Om mer än 55–60 procent återkommer och blir föremål för gastroskopi har vinsten med strategin försvunnit.

En annan faktor som har betydelse för kostnadsutfallet är i vilken utsträckning »opportunistisk» H pylori-screen-

ing (dvs i anslutning till läkarbesök av annan anledning) utförs även hos dem som initialt handläggs med prompt gastroskopi. Vidare förskjuts kostnaderna till »test and scope»-strategins nackdel om patienter med funktionell dyspepsi som vid den initiala H pyloritestningen befinns vara H pylori-positiva blir föremål för eradikeringsbehandling.

Ekonomiska synpunkter

Läkemedelsanvändningen för patienter med dyspepsi har i flera avseenden förändrats. En dramatisk ökning har skett av användningen av syrahämmande läkemedel, främst en mångdubbling av protonpumpshämmare. Detta beror på en hög förskrivning vid refluxsjukdom och funktionell dyspepsi. Den totala försäljningsvolymen protonpumpshämmare för gruppen med dyspepsi översteg år 1998 1,3 miljarder.

Enligt diagnos-receptundersökningen utgör ulkussjukdom ca 20 procent av alla fall med dyspepsi; motsvarande siffra för funktionell dyspepsi är ca 50 procent. Den diagnostiska träffsäkerheten är inte validerad. Med viss sannolikhet är andelen ulkuspatienter övervärderad; enligt många epidemiologiska studier torde den ligga närmare 10 procent. Insjuknandet har legat stabilt eller tenderat att sjunka på senare tid. Förskrivning av magsårsmedicin för funktionell dyspepsi kan således beräknas till ca 1/2 miljard kronor per år, vilket är anmärkningsvärt eftersom indikationen inte är godkänd och medlens effektivitet inte är vetenskapligt dokumenterad. Samhällets totala kostnader, direkta och indirekta, för dyspepsi kan uppskattas till 3,7–4,4 miljarder kronor per år.

En inträngande jämförelse av kostnadsutvecklingen över tiden är inte möjlig, då tidigare studier inte omfattat identiska diagnosgrupper. Vad som ändå framkommer är att kostnader för kirurgiska åtgärder sjunkit kraftigt, medan läkemedelskostnadernas andel ökat markant, från 2–3 procent till 25–30 procent av de totala.

Angelägen forskning

I och med upptäckten av sambandet mellan H pylori och ulkussjukdom restes frågan varför endast en minoritet av infektionsbärarna drabbas. Fördjupad kunskap krävs om de mekanismer som angriper respektive skyddar slemhinnan i magsäck och tolvfingertarm. Prospektiva epidemiologiska studier får ses som nödvändiga för att fastställa livsstils- och miljöfaktors roll. Redan tillgänglig kunskap om bakteriens arvs-massa kan förväntas ge information om olika bakteriestammars sjukdomsalstrande förmåga, information av betydelse

se för uppkomst inte enbart av ulkus utan även av magsäckscancer.

Funktionell dyspepsi, detta synbarligen heterogena tillstånd, medför mindre allvarliga symtom och följder än ulkus. Ur individens synvinkel är en kartläggning av de symtomgivande mekanismerna angelägen, eftersom nya farmakologiska vägar kan öppnas. Forskningen kan således leda fram till en rationell terapi av funktionell dyspepsi, något som uppenbarligen är av samhällsekonomiskt intresse.

Slutsatser

- Dyspepsi, ett samlingsbegrepp för smärta eller obehag lokaliserade centralt i övre delen av buken, är vanligt förekommande i befolkningen. Största informationsvärdet vid utredning av dyspepsi har undersökning med gastroscop. Antalet årliga undersökningar har ökat markant.

- Samhällets totala kostnader för utredning och behandling av patienter med dyspepsi är betydande. Med modern läkemedelsbehandling har de kostnader som är relaterade till kirurgiska operationer sjunkit markant under de senaste decennierna. I gengäld har kostnader för läkemedel stigit kraftigt. Dessa kostnader skulle kunna reduceras avsevärt om kunskapen om användningen av läkemedel vid de olika tillstånden som ingår i dyspepsi utnyttjades fullt ut i klinisk praxis.

- Mer än hälften av patienterna med dyspepsi får diagnosen funktionell dyspepsi, vilket innebär att någon orsak till besvären inte kunnat objektivt fastställas. Sannolikt är gruppen funktionell dyspepsi heterogen och består av flera undergrupper. Kunskapen om vad som framkallar och orsakar besvären vid funktionell dyspepsi är för närvarande begränsad. Det är angeläget att forskning bedrivs på detta område, så att effektiv och rationell behandling kan erbjudas.

- Symtomlindring vid funktionell dyspepsi kan uppnås med syrahämmare när besvären har ett dominerande inslag av reflux av syra till matstrupen. Prokinetiska läkemedel, vars verkan är att stimulera magsäckens rörelse, kan vara symtomlindrande i vissa fall. Det vetenskapliga underlaget är emellertid begränsat.

- Majoriteten av patienter med ulkussjukdom är kroniska bärare av bakterien *Helicobacter pylori*, som orsakar inflammatoriska förändringar i magsäcken. En mycket stor andel av befolkningen bär utan några symtom på denna infektion, men endast en av sex drabbas

av ulkussjukdom. Det är därför inte meningsfullt att införa massundersökningar där man söker efter bärare av infektion med *H pylori*. Kunskap om faktorer knutna till olika stammar av *H pylori*, liksom om individrelaterade faktorer, är angelägen forskning för att öka förståelsen för uppkomst av ulkussjukdom.

- Endast patienter med ulkussjukdom skall behandlas med antibiotika riktade mot bakterien *H pylori*. Det är vidare väl vetenskapligt dokumenterat att behandling med läkemedel som hämmar bildningen av saltsyra i magsäcken, speciellt av typ protonpumpshämmare, i kombination med två antibiotika inte bara läker såren i tolvfingertarmen utan framför allt minskar risken för återfall. Denna behandlingsstrategi är väl etablerad för sår både i tolvfingertarmen och i magsäcken, om än mindre väl underbyggd för magsäckssår.

- Patienter med refluxsjukdom, dvs med återkommande besvär på grund av uppstötning av maginnehåll genom övre magmunnen till matstrupen, blir snabbt symtomfria med saltsyrahämmande läkemedel av typ protonpumpshämmare.

- Läkemedel inom gruppen NSAID-preparat medför ofta biverkningar i form av dyspepsi och utgör en vanlig orsak till ulkus, framför allt i magsäcken. De används vid akuta och kroniska led- och muskelsjukdomar. Vid långvarig NSAID-medicinering, såsom för kronisk ledgångsreumatism, skall profylaktisk behandling övervägas hos patienter med riskfaktorer. Det är relativt väl underbyggt att läkemedlen protonpumpshämmare och misoprostol, med förmåga att både hämma saltsyrabildningen och skydda magsäckens slemhinna, är de mest verksamma vid sår-sjukdom i magsäcken, medan de syrahämmande medlen protonpumpshämmare och histamin-H₂-receptorblockerare är de mest verkningsfulla vid sår i tolvfingertarmen.

- Svåra och alarmerande magbesvär, exempelvis blödning, bör föranleda akut eller snar undersökning av matstrupe och magsäck med gastroscop. I vissa fall används undersökning av förekomst av *H pylori* för att ta ställning till om gastroscopi skall utföras eller ej. För- och nackdelar med olika strategier av handläggning belyses och diskuteras i rapporten med hjälp av en modell som tar upp olika beslutsvägar. Kliniska undersökningar med praktisk tillämpning av denna och likartade teoretiska beslutsmodeller vore av värde, såväl för den enskilde patienten som för samhällsresurserna.

- Magsäckscancer, som svarar för mindre än 1 procent av all dyspepsi, förekommer sällan under 45 års ålder. Med *H pylori*-infektion föreligger något förhöjd risk att drabbas av magsäckscancer. Det är för närvarande inte rationellt varken att utföra massundersökning av denna bakterieinfektion eller att behandla i cancerpreventivt syfte, eftersom flera andra faktorer är betydelsefulla. Vi saknar tillräcklig kunskap om skilda bakteriestammars virulens och om individens mottaglighet, vilket är en förutsättning för prevention.

Rapporten kan beställas från SBU, Box 5650, 114 86 Stockholm.

Tel 08-412 32 00.

Fax 08-411 32 60.

Se även medicinsk kommentar i detta nummer.

Leif Bjermer, professor, lunge sektionen, medisinsk avdelning, RIT, Trondheim (*leif.bjermer@medisin.ntnu.no*)

SBU-rapporten om astma och kroniskt obstruktiv lungsjukdom

Ett starkt stöd inför svåra beslut i den kliniska vardagen

II Obstruktiv lungsjukdom är ett ökande problem inte minst i västvärlden. Som det också påpekas i SBU-rapporten, som sammanfattas i detta nummer av *Läkartidningen*, betraktas idag både astma och kroniskt obstruktiv lungsjukdom (KOL) som folksjukdomar som beräknas kosta samhället ca 6 miljarder kronor varje år. Trenden under de senaste decennierna har varit att båda sjukdomsgrupperna i huvudsak sköts ambulantly, antingen hos specialist eller i primärvården. Tidig intervention, patientinformation samt adekvat medicinering är viktiga hörnpelare i behandlingen av dessa patienter.

Samtidigt som behandlingen har flyttats ut i primärvården har den paradoxalt nog blivit något mer komplicerad. Det ställs krav på en helhetsbild med behandling av »den sjuke patienten«. Som komplement till förskrivning av läkemedel måste vi idag både erbjuda undervisning och etablera ett partnerskap med patienten i avsikt att få den sjuke att själv, till största delen, ta ansvar för sin egen behandling. Målsättningen skall vara inte bara att behandla aktuell sjukdom utan också att förebygga, så att risken för återfall reduceras. Receptet är således utbytt mot ett behandlingspaket.

Situationen blir inte mindre komplicerad av att vi idag också har betydligt fler medikamenter att välja mellan. Man kan sammanfattningsvis säga att ett mer omfattande behandlingsutbud skapar en mer upplyst och, förhoppningsvis, en mer behandlingsmotiverad patient. För läkaren har rollen av undervisare blivit alltmer betonad.

Doktorn som lärare

Rollen som behandlare och undervisare ställer stora krav på att doktorn håller sig uppdaterad. Kunskaperna från de grundläggande läkarstudierna blir snabbt föråldrade, kontinuerlig vidareutbildning är ett måste. Den kunskap som skall inhämtas kan sammanfattas med:

- egen och andras kliniska erfarenhet,
- läroböcker,
- kurser och seminarier,
- egen litteraturgenomgång,
- »guidelines«, konsensusrapporter.

En explosion av artiklar om evidensbaserad medicin

Evidensbaserad medicin (EBM) är ett relativt nytt begrepp, som baserar sig på systematisk genomgång av studier publi-

Se även efterföljande medicinska kommentar samt artiklarna med start på sidan 288.

cerade i vetenskapliga tidskrifter. Man tar utgångspunkt i det faktum att endast ca 15–40 procent av medicinska beslut tas på grundval av befintliga vetenskapliga bevis [1]. Avsikten med EBM är att man genom en systematisk och kritisk litteraturgenomgång skall kunna fastställa vilken grad av dokumentation som ligger till grund för de beslut som tas i den kliniska vardagen.

Behovet av evidensbaserad medicin visar sig i den formliga explosion av artiklar som publicerats på temat under de senaste tio åren. I en sökning på Medline fann jag år 1990 endast två artiklar på temat EBM, medan motsvarande siffra för 1999 var 1 557 artiklar. Med sökord »asthma and treatment« eller »COPD and treatment« fann jag under perioden 1989–1999 16 008 artiklar om astma och 20 867 artiklar om »chronic obstructive pulmonary disease« (COPD).

Det finns således ett stort behov av assistans när det gäller att rätt söka i den dokumentation som finns (Figur 1).

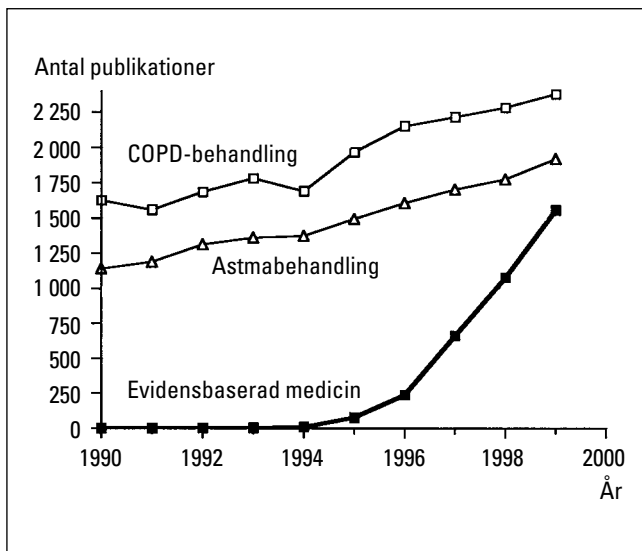
Graderingsprinciper

I SBU-rapporten om behandling av astma och KOL har ett stort antal av landets ledande experter inom området försökt ge en samlad översikt över den dokumentation som finns. Man har gått mycket systematiskt tillväga och utgått från litteratursökning i Medline från 1966–1999. I tillägg har man studerat både systematiska översikter från Cochrane-biblioteket och andra översikter som kvalitetsgranskats av Centre for Review and Dissemination vid universitetet i York (databasen DARE).

Som evalueringsparametrar har man valt:

- mortalitet,
- exacerbationer, attacker,
- hälsorelaterad livskvalitet,
- symtomskalor, symtompöäng.

Man har inkluderat enbart studier som pågått minst tre månader och som haft en randomiserad placebokontrollerad design. I tillägg skall studien som evalueringsvariabel ha minst ett av de fyra alternativen ovan. Studierna graderades vidare



Figur 1. En sökning i Medline ger stort utbyte. På senare år har antalet artiklar starkt ökat. Om man inte vill ha alltför många »träffar« måste sökkriterierna väljas med omsorg.

utifrån kvalitet och klinisk evidens och indelades sedan utifrån en fyrgradig kvalitets skala från 1 till 4, där 1 innebär stark dokumentation och 4 innebär obefintlig eller ingen dokumentation.

För- och nackdelar med EBM

SBU-rapportens författare redovisar på ett förtjänstfullt sätt hur man har gått tillväga. Möjligen skulle man ha kunnat vara lite mer kritisk vad avser EBM-metodens för- och nackdelar. Många av de studier som idag produceras inom astma- eller KOL-behandlingens område omfattar patienter som selekterats utifrån mycket strikta kriterier.

Ett exempel: Den idealiske astmastudiepatienten är icke-rökare, har en stabil instabil sjukdom, en basal FEV_{1s} mellan 60 och 85 procent av normalvärdet och en FEV₁-reversibilitet på 12, helst 15 procent. Kort sagt: En rökande astmatiker med diabetes och frekventa exacerbationer platsar inte i några läkemedelsstudier.

Detta innebär att vi ofta får extrapolera kunskap vi erhållit från strikt selekterade studiepatienter till de patienter vi möter i den kliniska vardagen. Evidensbaserad medicin skall därför behandlas för vad den är, dvs ett starkt stöd för beslutsfattning i den kliniska vardagen komplementärt till annan beprövad erfarenhet.

Rapporten är lättläst och logiskt uppställd

Uppställningen är logisk och gör att rapporten, trots omfånget, blir lättläst och fungerar som ett referensverk från första stund. I början av varje kapitel sammanfattas dess konklusioner, vilka graderas utifrån vilken bevisgrad som finns bakom påståendet. I slutet av varje kapitel redovisas i tabellform design och konklusion från de studier som tagits med i evalueringen. Man redovisar också vilken forskning och vilka studier som bör göras för att få fram bättre dokumentation i framtiden. På slutet finns en mer utförlig referenslista.

Nedan följer en del kommentarer till enskilda kapitel i rapporten.

Förebyggande icke-medikamentell behandling

Modern astma- eller KOL-behandling omfattar mer än den rent farmakologiska terapin. I rapporten finns en hel del nyt-

tig läsning på temat icke-farmakologisk behandling som spänner över teman som sekundär prevention, fysikalisk rehabilitering och psykologisk behandling. Man försöker även få fram dokumentation på effekt av alternativ terapi, en dokumentation som visar sig i stort sett saknas. Samma slutsatser drar man avseende effekt av psykologisk behandling av astma.

Sekundär prevention är ett svårt ämne, som författarna tacklar på ett förtjänstfullt sätt. Man slår bl a fast att det finns goda bevis för att luftrenare inte har någon plats i astmabehandlingen och att underlag saknas för att bedöma huruvida sekundärprevention har någon plats i KOL-behandlingen. Den viktigaste sekundärpreventionen är sannolikt att undvika såväl aktiv som passiv rökning. Rökproblemet har därför fått ett eget kapitel.

Lite överraskande fastslår man att rökstopp hos yngre och medelålders patienter med lindrig KOL inte påverkar dödligheten under en observationsperiod på åtta år (evidensgrad 1). Denna slutsats blir i viss mån en konsekvens av en systematisk litteraturgenomgång med mortalitet som en av evalueringsvariablerna. Man kan kanske inte förvänta en ökad dödlighet i den aktuella åldersgruppen. Om frågan i stället varit huruvida rökning hos yngre och medelålders patienter med lindrig KOL påverkar generell morbiditet under en observationsperiod på åtta år hade kanske svaret blivit något annorlunda.

En annan viktig upplysning är vad man kan förvänta sig när rökande astmatiker slutar röka. Man konkluderar att det saknas dokumentation på att rökstopp hos astmatiker kan bidra till remission, dvs symtomlindring eller minskat behov av astmamedicinering.

Klimatvård behandlas i ett separat kapitel. Även här finns det få studier. Författarna fann tre med tvekan acceptabla studier. Ingen hade kontrollmaterial, och evalueringsparametrarna var något dåligt definierade. Slutsatsen blev att klimatvård (läs: patientresa till utlandet) inte har någon dokumenterad effekt i behandlingen av astma eller KOL.

Farmakologisk behandling

Den farmakologiska behandlingen utgör naturligt nog en stor del av rapporten. Det är också här som man finner det största antalet publicerade artiklar. Detta ställer i sin tur stora krav på dem som utför en systematisk litteraturgenomgång och som måste gå igenom ett stort antal arbeten innan man kan selektera dem som uppfyller definierade kvalitetskrav.

Beta-2-agonister och antikolinergika. I kapitlet om beta-2-agonister och antikolinergika fastslår man, för en del något förvånande, att beta-2-agonister är dåligt dokumenterade vid behandling av akut astma hos barn (evidensgrad 3). Å andra sidan konkluderar man att användandet av racemiskt adrenalin som alternativ till beta-2-agonist inte har några fördelar hos barn med akut astma (evidensgrad 2). Däremot finner man att effekten hos vuxna har en väl dokumenterad effekt vid akut astma (grad 1) samt att kombinationen med antikolinergikum är gynnsam i denna situation (grad 1).

Denna skillnad mellan vuxna och barn återspeglar svårigheten att göra placebokontrollerade studier i akut fas på barn. Samtidigt ger den en varnande fingervisning om att det kanske inte alltid är rätt att direkt extrapolera behandlingsresultat från vuxna till barn.

Vad gäller långtidsbehandling med beta-2-agonist konkluderar man att regelbunden behandling med korttidsverkande beta-2-agonist är dåligt dokumenterad. Man har en klar tilläggseffekt av att addera långverkande beta-2-agonist till ste-

roider eller befintlig terapi hos vuxna (grad 1), medan dokumentationen är svag hos barn (grad 3). I det sistnämnda fallet kan man visa på ett antal välgjorda studier på barn som trots detta inte har kunnat visa på någon entydigt positiv effekt.

Steroider. Inhalationssteroidernas värde när det gäller att förebygga exacerbationer och eventuellt behov av sjukhusvård är väl dokumenterat hos såväl barn som vuxna (grad 1). Det finns viss dokumentation på att steroider även kan reducera risken för astmadödighet hos vuxna (grad 3), medan dokumentation helt saknas avseende barn (grad 4).

Kanske finner man, som förväntat, också dålig dokumentation på effekt av steroidbehandling hos små barn (yngre än 2 år).

När det gäller risk för systembiverkningar drar man slutsatsen att det finns visst stöd för att inhalationssteroider hämmar längdtillväxt (grad 2), kan ge skör hud med hudblödningar (grad 2) och kan öka risken för grå starr (grad 2).

När det gäller KOL konkluderar man att det är klart dokumenterat att steroider inte påverkar det naturliga förloppet (grad 1), och att det finns viss dokumentation på att inhalationssteroider kan reducera exacerbationsfrekvensen hos avancerad KOL (grad 3).

Annan »antiinflammatorisk« behandling. I kapitlet om kromoglikat drar man slutsatsen att det finns dokumenterad effekt hos vuxna och barn över 4 år (grad 1). Kombinationsbehandling med steroider nämns inte.

En annan del jag saknar är en kommentar kring kromoglikatbehandling vid ansträngningsutlöst astma. Detta är en indikation som man ofta finner i litteraturen, men där dokumentationen, framför allt jämförelsestudier mot annan terapi, är mager. Då mekanistiska variabler av typen »exercise induced bronchial obstruction« inte ingår som en effektvariabel i SBU-rapporten kommer dessvärre inte denna del med.

Anti-leukotriener. Behandling med leukotrienantagonister (LTRA) är en relativt ny behandlingsprincip mot astma. Man konkluderar att effekten är väl dokumenterad vid lindrig och moderat astma (grad 1) samt att det finns en dokumenterad tilläggs effekt hos både ungdom och vuxna vid samtidig behandling med inhalationssteroider (grad 1).

Underlaget är fortfarande bristfälligt vad avser behandling av små barn och patienter med KOL (grad 4). Även här saknar jag kommentarer kring behandling med LTRA på ansträngningsinducerad astma.

Annan behandling

Kapitlet om oxygenbehandling ger en översikt över både ny och väletablerad kunskap inom området. Oxygenbehandling förlänger överlevnaden hos KOL-patienter med kronisk hypoxi (grad 1) förutsatt att behandlingstiden är tillräckligt lång (grad 1).

Andra viktiga kliniska problemställningar som tas upp är behandling av patienter med ansträngningsinducerad hypoxemi samt värdet av oxygensköterska.

Det är uppenbart att det behövs fler studier inom området.

Volymreducerande kirurgi är en annan verksamhet som förtjänar en kritisk granskning. Emfysemkirurgi är resurskrävande; man bör förvissa sig om att behandlingen verkligen nyttar. Man hänvisar till en del pågående studier bl a i Sverige och att det finns indikation på, men fortfarande är något dåligt doku-

menterat, att behandlingen har positiv effekt på livskvalitet och överlevnad.

Ett viktigt referensverk

Rapporten ger utförliga svar på ett stort antal viktiga kliniska frågor om behandling av KOL och astma. Speciellt viktigt är att man inte bara fokuserat på den medicinska behandlingen utan också berör icke-farmakologisk terapi – såväl sekundär prevention som alternativ medicin och kirurgi. Rapportens utformning gör den lättläst och väl ägnad att fungera som referensverk.

Det bör noteras att man i den kritiska och systematiska genomgången tar upp några viktiga, men specifika, frågeställningar som vanligtvis besvaras utifrån ett starkt selekterat patientmaterial. I den kliniska vardagen är det dock viktigt att tillgänglig information, som alltid, kombineras med klinikers erfarenhet och förnuft.

Referens

1. Mulligan JE, Rowe J, Sacket DL. Inpatient general medicine is evidence based. *Lancet* 1995; 346: 407-10.

Syrgasbehandling oomtvistad vid svår, men tveksam vid medelsvår, hypoxi

Kommentar till metaanalys av hemoxygen vid kroniskt obstruktiv lungsjukdom

|| Vid kroniskt obstruktiv lungsjukdom (KOL) kan förloppet kompliceras av hypoxi i samband med exacerbationer, ansträngning eller sömn. Kronisk hypoxi i vila dagtid trots optimerad behandling tillstöter hos en del patienter med avancerad sjukdom och denna hypoxi är förknippad med starkt förkortad överlevnad [1] och nedsatt hälsorelaterad livskvalitet [2]. Överlevnaden för patienter med KOL och kronisk hypoxi utan oxygenbehandling är betydligt sämre än överlevnaden vid motsvarande grad av lungfunktionsnedsättning utan hypoxi. Den kroniska hypoxin har en negativ inverkan på en rad fysiologiska parametrar som intellektuell förmåga, njurfunktion, endokrina funktioner och muskelmetabolism [3, 4]. Kronisk hypoxi kan obehandlad leda till polycytemi liksom till pulmonell hypertension med risk för ytterligare komplikationer.

Avsikten med långtidsbehandling med oxygen i hemmet är att eliminera den överdödlighet hypoxin orsakar, förbättra livskvaliteten och att eliminera hypoxiorsakad funktionsnedsättning och symtom. Det finns också förhoppning om att behandlingen skall minska sjukvårdsbehovet.

Denna översikt kommenterar en systematisk litteraturgranskning från Cochrane collaboration angående hemoxygenbehandling vid KOL, vad gäller dess effekter på överlevnad, livskvalitet och förbättring av fysiologiska parametrar [5].

Litteraturgranskningens material och metoder

Granskningen omfattar randomiserade, kontrollerade studier som jämförde resultaten för patienter med hypoxi och KOL som antingen erhöll långtidsbehandling med oxygen, hemoxygenbehandling eller ingick i en kontrollgrupp. Studierna identifierades ur Cochrane airways group COPDs (chronic obstructive pulmonary disease) register med hjälp av sökorden hem(behandling) och oxygen. Vidare sökte man bland referenserna i de identifierade artiklarna och kontaktade företag som levererar oxygenutrustning och medlemmar i International respiratory care club angående opublicerade studier. Man kontaktade författarna till de publicerade studierna för att få ytterligare data för metaanalys, dock utan framgång.

Litteraturgranskningen inkluderade studier på vuxna patienter med KOL som erhöll långtidsbehandling med oxygen utanför sjukhus. Majoriteten av patienterna hade kronisk hypoxi med $\text{PaO}_2 < 7,3$ kPa, men en del hade $\text{PaO}_2 > 7,3$ kPa i vila med antingen nattlig hypoxi eller hypoxi vid ansträngning. Alla typer av oxygenutrustning användes (koncentrator, fly-

SAMMANFATTAT

Cochrane collaboration har genomfört en litteraturgranskning av hemoxygenbehandling vid kronisk obstruktiv lungsjukdom. Effekten på överlevnad, livskvalitet och fysiologiska parametrar bedömdes.

Följande slutsatser drogs: hemoxygenbehandling förbättrar överlevnaden vid svår kronisk hypoxi ($\text{PaO}_2 < 8$ kPa), men inte vid lätt till medelsvår hypoxi; oxygenbehandling nattetid har ingen effekt på överlevnaden vid nattlig hypoxi utan svår hypoxi dagtid; statistiskt signifikanta förbättringar i en del fysiologiska variabler har liten mätbar effekt på överlevnad och livskvalitet.

Slutsatsen att överlevnadsvinst erhålls vid arteriell $\text{O}_2 < 7,4$ kPa i vila dagtid är odisputabel. En studie visade ingen överlevnadsvinst vid arteriellt $\text{O}_2 7,4-8$ kPa. Vid denna hypoxigrad är överlevnadsvinsten alltså synnerligen tveksam.

Evidensbaserad medicin

tande oxygen, gasflaskor) och kontrollgruppen erhöll antingen luft via samma utrustning, oxygen under kortare tid av dygnet eller ingen oxygenbehandling.

Man planerade göra följande subgruppsanalyser i en metaanalys:

- jämförelse mellan män och kvinnor;
- kontinuerlig oxygenbehandling jämfört med nattlig behandling;
- oxygenbehandling jämfört med ingen behandling;
- nattlig oxygenbehandling jämfört med rumsluft vid hypoxi under sömn;
- oxygenbehandling jämfört med ingen behandling vid måttlig hypoxi.

Litteraturgranskningens resultat

Man identifierade sex artiklar angående fyra studier, varav en artikel exkluderades då den enbart innehöll data vid studie-

starten. Data från studierna kunde inte aggregeras till någon metaanalys på grund av skillnader i inkluderade patienter och behandling. Resultaten i de fyra studierna beskrevs.

Kontinuerlig oxygenbehandling jämfört med nattlig oxygenbehandling (NOTT-studien) [6-8]: 1 043 patienter från sex kliniker i USA undersöktes för eventuell inklusion, varav 809 exkluderades av olika orsaker, exempelvis förbättring av PaO₂ till över 7,9 kPa vid uppföljning efter 1–3 veckor. Patienterna hade alla PaO₂ < 7,9 i vila dagtid, och de patienter som hade PaO₂ i intervallet 7,4–7,9 kPa hade antingen polycytemi, ödem eller tecken på hjärtpåverkan. De 203 inkluderade patienterna randomiserades till kontinuerlig oxygenbehandling eller nattlig oxygenbehandling. Oxygenutrustningen var koncentrator, flytande oxygen eller gasflaskor. I gruppen som erhöll kontinuerlig behandling var 77 procent män och medelåldern 65 år, medan 80 procent var män och medelåldern 66 år i gruppen som erhöll nattlig behandling.

Ettårsöverlevnaden skilde sig inte mellan grupperna, men vid två år var överlevnaden signifikant bättre i gruppen som erhöll kontinuerlig behandling (oddskvot för död 0,45; 95 procents konfidensintervall 0,25–0,81) [6].

Vidare fann man små förbättringar i livskvalitet och intellektuell förmåga när data från alla patienter i studien kombinerades [7]. Man fann också bättre effekt på trycket i arteria pulmonalis, pulmonellt vaskulärt motstånd och hjärtats slagvolym i den patientgrupp som fått kontinuerlig oxygenbehandling [8].

Oxygenbehandling jämfört med ingen oxygenbehandling (MRC-studien) [9]: Vid tre kliniker i Storbritannien randomiserades 87 patienter med PaO₂ i intervallet 5,3–8 kPa i vila dagtid om de hade haft en eller flera exacerbationer med underbensödem till behandling med oxygen 15 timmar/dygn eller ingen oxygenbehandling. Alla hade eller hade haft koldioxidretention, och de utgjorde alltså en starkt selekterad patientgrupp. Liksom i NOTT-studien användes alla typer av oxygenutrustning. Medelåldern i de olika patientgrupperna låg mellan 56 och 59 år. Analyserna redovisas separat för män (33 behandlade, 33 kontroller) och kvinnor (9 behandlade, 12 kontroller).

Efter fem år noterades en signifikant bättre överlevnad i de behandlade grupperna (oddskvot för död 0,42; 95 procents konfidensintervall 0,18–0,98). Förbättringen i överlevnad sågs först efter de 500 första behandlingsdagarna för männen, men från början av studien för kvinnorna. Antalet kvinnor var dock litet.

Hög summa av värdena för PaCO₂ (mm Hg) plus röd blodkroppsmassa (ml/kg kroppsvikt) var den bästa prediktorn för tidig död både bland behandlade och kontroller, och en summa > 98 indikerade sannolikt sämre effekt av långtidsbehandling med oxygen.

Nattlig oxygenbehandling jämfört med luft (Fletcher) [10]: I denna dubbelblinda studie randomiserades 38 patienter med nattlig hypoxi och PaO₂ dagtid > 8 kPa till behandling med oxygen eller luft från koncentrator. Kontrollpatienterna fick dock maximalt 25 procent oxygen i stället för rumsluft (med 21 procent oxygen) från sina koncentratorer. Ingen skillnad i överlevnad eller antalet patienter som utvecklade kronisk hypoxi dagtid, och alltså behövde kontinuerlig oxygenbehandling, sågs.

Eftersom studien omfattade få patienter kan man inte utesluta typ 2-fel. Dock var oddskvoten för död mycket nära 1.

Långtidsoxygenbehandling jämfört med ingen oxygenbehandling vid måttlig hypoxi (Gorecka) [11]: Vid nio regionala kliniker i Polen randomiserades 135 patienter med KOL och måttlig hypoxi, PaO₂ i intervallet 7,4–8,7 kPa till behandling med

oxygenkoncentrator eller ingen oxygenbehandling. Medelåldern var 60 år i behandlingsgruppen och 62 år i kontrollgruppen. Ingen skillnad i överlevnad under den treåriga uppföljningstiden sågs. Oddskvot för död för de oxygenbehandlade patienterna var 1,09; 95 procents konfidensintervall 0,68–1,75.

Litteraturgranskningens diskussion och konklusioner

MRC-studien och Goreckas studie jämförde båda oxygenbehandling med ingen oxygenbehandling, men någon metaanalys gjordes inte eftersom studierna inkluderade patienter med olika grad av hypoxi. Kontrollgruppernas behandling varierade från ingen till nattlig oxygenbehandling respektive 25 procent oxygen i luft från koncentrator. I NOTT-studien anges inte hur många patienter som fick de olika typerna av oxygenbehandling. Endast Fletchers studie var dubbelblind, eftersom det inte går att ge flytande oxygen blint.

Två av studierna, NOTT- och MRC-studierna, rapporterade signifikant förbättrad överlevnad. Endast MRC-studien redovisade resultatet separat för män och kvinnor, och här sågs en överlevnadsvinst från studiestarten bland kvinnorna medan det dröjde 500 dagar innan någon överlevnadsvinst sågs bland männen. Patienterna i MRC-studien hade högre PaCO₂ vid studiestarten, och granskarna kommenterar att detta inte tycks ha påverkat resultaten.

För MRC-studien var »number needed to treat« (1/absolut relativ risk) 5, vilket betyder att under den femåriga behandlingstiden vann man ett liv per fem patienter behandlade med oxygen.

Nattlig oxygenbehandling för patienter med nattlig hypoxi under sömn men PaO₂ dagtid > 8 kPa förefaller inte ge någon överlevnadsvinst. Man kan inte dra någon säker slutsats på denna punkt, bland annat eftersom patientantalet var så litet i Fletchers studie. Att kontrollgruppen kan ha fått 25 procent oxygen i stället för luft kan ha lett till en höjning av PaO₂ för kontrollpatienterna och därmed vara en bidragande förklaring till den observerade avsaknaden av effekt.

Patienterna i de utförda studierna är inte representativa för de KOL-patienter som ordinerar långtidsbehandling med oxygen i hemmet idag. Dagens patienter är äldre och har flera andra sjukdomar utöver KOL. Patienter med komplicerande sjukdomar uteslöts ur de randomiserade, kontrollerade studierna. Man kan alltså inte dra slutsatsen att dagens patienter, som uppfyller riktlinjerna för långtidsbehandling med oxygen verkligen har nytta av behandlingen. Patienternas rökvanor anges varken i NOTT- eller MRC-studien, och inte heller finns någon analys som visar om rökvanorna påverkat överlevnaden. Avsaknaden av data om rökvanor gör också att vi kanske inte kan överföra resultaten till dagens situation då vi i praktiken kräver rökstopp före start av långtidsbehandling med oxygen i hemmet.

Betydelsen av »body mass index« diskuteras inte i någondera studien. Vidare saknas uppgifter om exacerbationsfrekvenser i samtliga artiklar.

Granskarna kommenterar att det är möjligt att statistiskt signifikanta förbättringar i en del fysiologiska variabler har liten mätbar effekt på patientens upplevda livskvalitet eller överlevnad.

Sammanfattningsvis kommer granskarna fram till att långtidsbehandling med oxygen i hemmet förbättrar överlevnaden i en selekterad grupp KOL-patienter med svår hypoxi. Däremot tycks behandlingen inte ha någon effekt på överlevnaden vid måttlig hypoxi eller nattlig hypoxi utan svår hypoxi dagtid. Vad gäller långtidsbehandling med oxygen till patienter med måttlig hypoxi och isolerad nattlig hypoxi behöver frågan studeras ytterligare i studier med större patientantal. Av etiska skäl tvekar man att randomisera patienter med svår hypoxi till placebo, och detta kan vara ett hinder för nya stu-

dier av effekten av långtidsbehandling med oxygen vid svår hypoxi.

Min kommentar

Granskarna har inte klart definierat vilka indikationer för hembehandling med oxygen vid KOL som skulle granskas. Jag saknar en granskning av behandling med oxygen vid ansträngningsorsakad hypoxi eller ansträngningsutlöst dyspné. Det finns en randomiserad kontrollerad dubbelblind studie med »cross over-design« av oxygen vid ansträngning som inte visade någon effekt av oxygen jämfört med luft på livskvaliteten under en sexveckorsperiod [12]. Hälften av patienterna föredrog luft, hälften placebo.

Litteraturgranskningen utfördes med hjälp av de artiklar som fanns i Cochrane airways group COPDs register, men utan någon regelrätt Medlinesökning. I stället granskades referenserna i de i registret funna sex artiklarna, och medlemmar i International respiratory care club tillfrågades. Detta ledde till att man missade en referens angående effekten av oxygenbehandling på intellektuell kapacitet och livskvalitet från NOTT-studien [7]. Detta har haft mindre effekt på granskarnas slutsatser. Man missade också en svensk randomiserad kontrollerad studie av effekten av behandling med flytande oxygen kontra koncentrator och konventionell gasflaska hos patienter med kronisk hypoxi [13]. Den studien visade signifikant bättre utveckling av livskvaliteten under en sexmånadersperiod för de KOL-patienter som fick flytande oxygen. Förbättringen iaktogs endast hos patienter med svår hypoxi ($\text{PaO}_2 < 7,4$ kPa). Resultatet talar för att det vore av värde att undersöka oxygenutrustningens betydelse för livskvaliteten hos patienter som inte är bundna till hemmet.

Den stora svagheten i översikten är att man inte har dragit någon slutsats om oxygenbehandlingens värde vid hypoxi med PaO_2 i gränsoområdet 7,4–8 kPa. Av översiktens abstract framgår att man med svår hypoxi avser patienter med $\text{PaO}_2 < 8$ kPa. Patienter med PaO_2 i detta gränsoområde ingick förvisso i båda de studier som visade överlevnadsvinst, men medelvärdena för PaO_2 här var 6,6–6,9 kPa, och majoriteten av patienterna hade alltså lägre PaO_2 . Det är då svårt att helt bortse från det resultatet i Goreckas studie, där 74 patienter hade ett PaO_2 i gränsoområdet 7,4–8 kPa med samma överlevnad som patienter med högre PaO_2 , alltså utan någon effekt av långtidsbehandling med oxygen i hemmet. Effekten på överlevnaden av långtidsbehandling med oxygen för patienter med hypoxi i intervallet 7,4–8 kPa måste alltså bedömas som osäker eller ingen, och dessa patienter bör inte omfattas av rekommendationer för långtidsbehandling med oxygen på bas av nuvarande kunskap. På denna punkt är alltså Cochrane-rapporten missvisande.

Effekten av långtidsbehandling med oxygen i hemmet på livskvalitet och fysiologiska parametrar är obetydlig, om någon, och kan knappast ligga till grund för rekommendation om behandling annat än i särskilda fall. Detta kan gälla patienter med polycytemi orsakad av hypoxi, där långtidsbehandling med oxygen visats ha en god effekt på hematokriten (NOTT) eller patienter med mer uttalad pulmonell hypertension. Kronisk hypoxi har effekt på en lång rad fysiologiska funktioner och granskarnas kommentar att effekten av oxygenbehandling på dessa funktioner ofta saknar effekt på patientens livskvalitet eller överlevnad är välmotiverad.

Under 1999 publicerades en av granskarna efterlyst studie av värdet av nattlig oxygenbehandling vid nattlig hypoxi [14]. Denna studie inkluderade 76 patienter med nattlig hypoxi och PaO_2 dagtid i intervallet 7,4–9,2 kPa, vilka randomiserades till nattlig oxygenbehandling eller ingen oxygenbehandling och följdes i medeltal 35 månader. Studien bekräftar helt Fletchers fynd: ingen effekt på överlevnaden och ingen effekt på antalet fall av svår hypoxi dagtid av nattlig oxygenbehandling.

Ytterligare en brist i de studier som publicerats är att pati-

enterna i så stor utsträckning är män. Endast i MRC-studien redovisades separata analyser för män och de få kvinnorna. Nu är majoriteten av de patienter som startar långtidsbehandling med oxygen i hemmet kvinnor [15]. De är yngre än männen vid behandlingsstarten och lever längre, varför de dominerar än mer i populationen patienter med långtidsbehandling.

Granskarna skriver att det högre medelvärdet för PaCO_2 hos patienterna i MRC-studien jämfört med NOTT-studien förefaller sakna betydelse. Dock var mortaliteten för de oxygenbehandlade männen i MRC-studien högre än i NOTT-studien under de första 500 behandlingsdagarna, och hög summa av PaCO_2 plus röd blodkroppsmassa befanns vara en indikator för död i MRC-studien. Selektionen av patienter med mer avancerad sjukdom till MRC-studien är alltså en möjlig orsak till den sämre överlevnaden – även andra studier visar ökad mortalitet bland patienter med högre PaCO_2 [16].

Värdet av hemoxygenbehandling vid svår kronisk hypoxi är oomtvistat. I Sverige har en femtedel av de cirka tusen patienter som varje år startar hemoxygenbehandling hypoxi av medelsvår grad [15]. En oreflektad inställning till behandlingsindikationen i dessa fall, där värdet av behandlingen varken visats vad gäller livskvalitet eller överlevnad, riskerar leda till meningslös behandling med dålig patientmedverkan. Vi behöver utvärdera behandlingens vinst för de enskilda patienterna när det gäller eventuell långtidsoxygenbehandling av medelsvår hypoxi.

Referenser

1. Renzetti AD Jr, McClement JH, Litt BD. The Veterans administration cooperative study of pulmonary function. 3. Mortality in relation to respiratory function in chronic obstructive pulmonary disease. *Am J Med* 1966; 41: 115-29.
2. Jones PW, Baeyestock CM, Littlejohns P. Relationship between general health measured with the sickness impact profile and respiratory symptoms, physiological measures and mood in patients with chronic airflow limitation. *American Review of Respiratory Disease* 1989; 140: 1538-43.
3. Grant I, Heaton RK, McSweeney AJ, Adams KM, Timms RM. Neuropsychologic findings in hypoxemic chronic obstructive pulmonary disease. *Arch Intern Med* 1982; 142: 1470-6.
4. Reihman DH, Farber MO, Weinberger MH, Henry DP, Fineberg NS, Dowdeswell IRG et al. Effect of hypoxemia on sodium and water excretion in chronic obstructive lung disease. *Am J Med* 1985; 78: 87-94.
5. Crockett AJ, Moss RJ, Cranston JM, Alpers JH. Domiciliary oxygen for chronic obstructive pulmonary disease (Cochrane review). In: *The Cochrane library* 1999. Issue 4. Oxford: Update Software, 1999.
6. Nocturnal oxygen therapy trial group. Continuous or nocturnal oxygen therapy in hypoxemic chronic obstructive lung disease. *Ann Intern Med* 1980; 93: 391-8.
7. Heaton RK, Grant I, McSweeney AJ, Adams KM, Petty TL. Psychologic effects of continuous and nocturnal oxygen therapy in hypoxemic chronic obstructive pulmonary disease. *Arch Intern Med* 1983; 143: 1941-7.
8. Timms RM, Khaja FU, Williams GW. Nocturnal oxygen therapy trial group. Hemodynamic responses to oxygen therapy in chronic obstructive pulmonary disease. *Ann Intern Med* 1985; 102: 29-36.
9. Medical research council working party. Long-term domiciliary oxygen therapy in chronic hypoxic cor pulmonale complicating chronic bronchitis and emphysema. *Lancet* 1981; 1: 681-6.
10. Fletcher E, Luckett R, Goodnight-White S, Miller CC, Qian W, Costaragos-Galarza C. A double-blind trial of nocturnal supplemental oxygen for sleep desaturation in patients with chronic obstructive pulmonary disease and a daytime PaO_2 above 60 mm Hg. *American Review of Respiratory Disease* 1992; 145: 1070-6.
11. Görecka D, Gorzelak K, Sliwinski P, Tobiasz M, Zielinski J. Effect of long-term oxygen therapy on survival in patients with chronic obstructive pulmonary disease with moderate hypoxaemia. *Thorax* 1997; 52: 674-9.
12. McDonald CF, Blyth CM, Lazarus MD, Marschner I, Barter CE. Exertional oxygen of limited benefit in patients with chronic ob-

- structive pulmonary disease and mild hypoxemia. *Am J Respir Crit Care Med* 1995; 152: 1616-9.
13. Andersson A, Ström K, Brodin H, Alton M, Boman G, Jakobsson P et al. Domiciliary liquid oxygen versus concentrator treatment in chronic hypoxaemia: a cost-utility analysis. *Eur Respir J* 1998; 12: 1284-9.
 14. Chaouat A, Weitzenblum E, Kessler R, Charpentier C, Ehrhart M, Schott R et al. A randomized trial of nocturnal oxygen therapy in chronic obstructive pulmonary disease patients. *Eur Respir J* 1999; 14: 1002-8.
 15. Oxygenbehandling och annan hemsjukvård. I: SBU. Behandling av astma och KOL. Stockholm: SBU, 2000. Rapport 151.
 16. Ström K. Survival of patients with chronic obstructive pulmonary disease receiving long-term domiciliary oxygen therapy. *American Review of Respiratory Disease* 1993; 147: 585-91.

SUMMARY

Oxygen therapy undisputed in severe, but doubtful in moderate, hypoxemia; comment to meta-analysis of domiciliary oxygen therapy in chronic obstructive pulmonary disease

Kerstin Ström

Läkartidningen 2001; 98: 295-8.

The review included randomized controlled trials on patients with chronic obstructive pulmonary disease (COPD) receiving domiciliary long-term oxygen therapy (LTOT). The authors identified six articles concerning four randomized controlled trials but could not perform any meta-analysis due to the heterogenous patient populations and treatments. From these trials they conclude that LTOT improves survival in patients with severe hypoxemia (arterial PaO₂ less than 8 kPa) but has no effect in patients desaturating only at night or in patients with moderate hypoxemia. They also sensibly remark that it is possible that statistically significant improvements in some physiological variables have little measurable impact on subjects perceived quality of life or survival.

The conclusion that survival benefit is demonstrated also in patients with hypoxemia in the range 7,4-8 kPa is debatable, with the strongest evidence pointing against benefit. These patients are better classified as moderately hypoxemic. In Sweden, they comprise 20% of new patients starting on LTOT. For them, the effect of LTOT should be evaluated individually in terms other than survival or quality of life.

A more recently published trial supports the conclusion that domiciliary nocturnal oxygen therapy has no impact on survival in nocturnal desaturation without severe daytime hypoxemia. There is also new evidence that the type of oxygen equipment might have a decisive impact on the quality of life in mobile patients receiving LTOT – improved quality of life with liquid oxygen and poorer quality of life with concentrator and conventional (heavy) gas cylinder. The importance of optimum equipment selection for each patient has been overlooked but merits further investigation

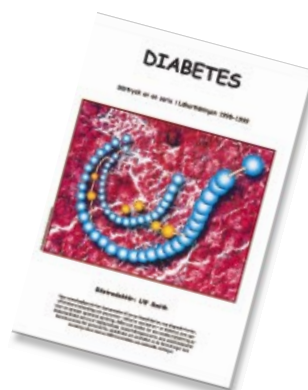
Correspondence: Kerstin Ström, Lung-/allergisektionen, Blekingesjukhuset, SE-371 85 Karlskrona, Sweden

Särtryck

Läkartidningen

Nya vetenskapliga rön har lagt grunden till en ny klassifikation, nya diagnoskriterier, effektivare behandling och prevention – inklusive vaccination – av diabetes, som uppvisar en närmast epidemisk spridning. Målen och medlen för den snabba förbättringen av diabetesvården som nu är möjlig belyses i Läkartidningens serie, som också analyserar konsekvenserna för patienterna, sjukvården och samhället av de förändringar som Socialstyrelsen hösten 1999 fastställde som nationella riktlinjer.

Priset är 85 kronor.



Diabetes

Beställer härmed.....ex
av "Diabetes"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till Läkartidningen
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Björn Stållberg, distriktsläkare, vårdcentralen, Trosa (bjorn.stallberg@nln.dll.se)

Bertil Marklund, docent, allmänläkare, FoU-enheten, primärvården Halland (bertil.marklund@lthalland.se)

SBU-rapporten om astma och KOL

Matnyttig och lättläst även för en jäktad allmänläkare

II Det är en imponerande systematisk litteraturgenomgång av behandling av astma och kroniskt obstruktiv lungsjukdom (KOL) som gjorts av SBUs projektgrupp [1]. Vad kan då en evidensbaserad kunskapssammanställning som SBU-rapporten få för betydelse för behandlingen av astma och KOL ur ett allmänmedicinskt perspektiv? Även om många följer de behandlingsrekommendationer som tidigare getts i olika sammanhang bör slutsatserna i rapporten få alla som behandlar denna patientgrupp att dels vara mer aktiva i behandlingen där det finns ett vetenskapligt stöd för detta, dels ompröva och ifrågasätta behandlingar som är mindre väl underbyggda.

Vid SBUs genomgång av befintlig litteratur har patienten satts i fokus. Man har vid bedömningen av olika studier använt sådana effektmått som är viktiga och betydelsefulla för patienten, såsom behov av ökad medicinering och akutbesök, hälsorelaterad livskvalitet, symtomskalor och dödlighet. Detta innebär att man vid värderingen av olika behandlingar i första hand sett till det som är viktigt för patienten, inte till måtvärden som t ex olika mått på lungfunktionen.

Givetvis behöver vi i den kliniska vardagen använda olika lungfunktionsmått vid diagnostik och bedömning av svårighetsgraden av astma och KOL. Bara anamnesen är otillräcklig för en korrekt diagnostik av astma respektive KOL. Vid KOL är också mätning av lungfunktionen nödvändig för att kunna bedöma såväl prognos som behov av andra resurser. Även om det är en dålig korrelation mellan lungfunktion mätt med FEV₁ och upplevd livskvalitet har man i det dagliga arbetet i många fall nytta av spirometri och/eller mätning av »peak expiratory flow« (PEF) vid uppföljning av patienter med astma.

Astma och KOL vanligt i primärvården

Att astma och KOL är mycket vanligt förekommande är väl känt från många epidemiologiska studier. Astma ökar främst bland barn och unga vuxna. Man beräknar att idag 6–8 procent av befolkningen har astma.

KOL ökar bland rökare, framför allt hos kvinnor. Av hela befolkningen över 50 års ålder har ca 8 procent KOL. I gruppen rökare som uppnår hög ålder utvecklar ungefär hälften KOL. Det är därför naturligt att båda dessa patientgrupper är vanliga i primärvården.

Se även föregående medicinska kommentar samt artiklarna med start på sidan 288.

I rapporten slås fast att KOL är en underdiagnostiserad sjukdom med såväl ökad prevalens som ökad morbiditet och mortalitet. Sjukdomen upptäcks i många fall alltför sent. För att ställa en korrekt och tidig diagnos vid KOL krävs spirometri.

Fråga patienten om rökning

Eftersom det är i primärvården som dessa patienter vanligtvis har sina flesta sjukvårdskontakter är det nödvändigt att resurser och möjligheter ges för tidig diagnostik av KOL. Rökstopp är här den viktigaste åtgärden. Spirometri bör därför göras betydligt oftare än vad som är fallet idag på rökare som söker i primärvården för olika luftvägssymtom. SBU-rapporten visar att tobaksrökning är den dominerande orsaken till KOL. Att alltid fråga om tobaksvanor, oberoende av var patienten söker, är enkelt och har visats ha effekt när det gäller rökstopp.

Rökstopp är den viktigaste behandlingen för patienter med KOL. Det finns dessutom ett vetenskapligt stöd för att rökstopp hos patienter med måttlig till svår KOL leder till en påtagligt förlängd överlevnadstid. Tyvärr saknas idag på många håll möjligheter att kunna erbjuda patienterna ett aktivt rökavvänningsstöd.

Fysisk träning och rehabiliteringsprogram

En av rapportens slutsatser är att fysisk träning är till nytta för patienter med medelsvår KOL. Den fysiska träningen leder till bättre livskvalitet, mindre andnöd och bättre arbetsförmåga. Förhoppningsvis kommer detta att leda till att fler patienter får tillgång till rehabiliteringsprogram där sjukgymnast, arbetsterapeut och i vissa fall även dietist ingår i behandlingsteamet. Utvärderingar behöver sedan göras för att bedöma hur dessa rehabiliteringsprogram kan utformas på bästa sätt.

Astmamottagning ett forskningsområde

Det påpekas i rapporten att astmamottagningar i primärvården kan ha en gynnsam effekt på symtom, akutbesök och

sjukskrivningar men att det vetenskapliga underlaget är bristfälligt. En slutsats i rapporten är att behovet av utvärderingar av astmamottagningar är stort.

Eftersom det idag är många som är övertygade om de positiva effekterna av astmamottagningar är det viktigt att primärvården även prioriterar forskning inom detta område. Framtida utvärderingar av astmamottagningar behöver belysa både patientnyttan – i form av bättre livskvalitet, mindre symtom, färre akuta exacerbationer – och kostnadseffektiviteten.

Inget stöd för alternativmedicinska behandlingsmetoder

Vi tillfrågas ofta av patienterna om så kända alternativa behandlingar kan ha någon effekt på deras sjukdom. I den litteraturgenomgång som gjorts saknas underlag för att dessa behandlingar, bl a akupunktur och homeopati, skulle ha någon effekt vid astma eller KOL.

Detta innebär att det inte finns något vetenskapligt stöd för att rekommendera alternativa behandlingsmetoder till dessa patienter.

Inhalationssteroider vid astma väldokumenterade

Det finns idag ingen tveksamhet till att inhalationssteroider utgör basbehandlingen vid astma. Detta bekräftas också i SBU:s kunskapssammanställning. Dessa läkemedel har också inneburit att kostnaderna för sjukhusvård av astmapatienter har minskat dramatiskt. Trots att inhalationssteroider är väldokumenterade och har mycket god effekt vid behandling av astma är den kliniska erfarenheten att många patienter med astma fortfarande är underbehandlade. Detta beror i många fall på dålig följsamhet till ordination från patientens sida, men därutöver kan säkert behandlingen ibland optimeras.

Det har ibland uttryckts en oro för biverkningar, bl a i form av osteoporos. Här slår rapporten fast att det inte finns några studier som skulle ge belägg för att behandling med inhalationssteroider i rekommenderade doser skulle ge upphov till osteoporos.

När det gäller akutbehandling redovisar rapporten att det finns ett starkt stöd för att använda peroral korttidskurer av steroider vid behandling av akuta exacerbationer av astma hos vuxna. Detta görs inte alltid, trots att det är rekommenderat sedan länge.

Vilka KOL-patienter skall ha inhalationssteroider?

Den kliniska effekten av underhållsbehandling med inhalationssteroider vid KOL är däremot inte lika uppenbar. Många patienter med KOL behandlas idag med inhalationssteroider, ibland alltför slentrianmässigt eller beroende på att diagnosen felaktigt ställts som astma.

Enligt SBU-rapporten finns det idag inga belägg för att inhalationssteroider långsiktigt skulle kunna förhindra den fortsatta progressen av KOL-sjukdomen. Däremot kan inhalationssteroider minska antalet exacerbationer vid avancerad KOL, vilket i vissa fall kan vara en indikation för användande av inhalationssteroider till denna patientgrupp.

Användning av beta-2-stimulerare

Vid underhållsbehandling av astma hos vuxna finns idag ett starkt stöd för att tillägg av långverkande beta-2-stimulerare till behandling med inhalationssteroider har god effekt enligt de effektmått som används i SBU-rapporten. Detta är också något som under de senaste åren blivit alltmer etablerat i klinisk praxis.

Även vid KOL finns underlag för att långverkande beta-2-stimulerare, liksom antikolinergika, har effekt, även om SBU konkluderar att de ger små förbättringar av symtom och livskvalitet. Hittills har inte långverkande beta-2-stimulerare

förordats vid KOL, men de behandlingsrekommendationer som idag finns kan kanske komma att ändras i framtiden. Det är inte ovanligt att det förskrivs mixturer med beta-2-stimulerare till barn under 1 år med hosta. SBU:s slutsats är att detta inte har någon påtaglig effekt, varför behandlingen med dessa mixturer bör kunna minskas vid obstruktiva tillstånd.

Det är anmärkningsvärt att placebokontrollerade studier saknas som bekräftar den kliniska erfarenhet som alla har avseende den gynnsamma effekten av beta-2-stimulerare vid behandling av akut astma.

Antileukotriener och teofyllamin

SBU-rapporten skriver att antileukotriener har en gynnsam effekt vid underhållsbehandling av lindrig till medelsvår astma. En viktig fråga som inte besvaras är var på behandlingsstappen som dessa läkemedel skall placeras. För att få svar på denna fråga behövs fler studier.

Rapporten slår fast att teofyllaminpreparat som underhållsbehandling vid astma inte har effekt. Dessutom har de en mycket ogynnsam biverkningsprofil. Användningen av teofyllamin har idag en mycket liten plats vid behandling av astma i primärvården. Användningen av peroral teofyllaminpreparat bör därför omprövas. Däremot, vilket är viktigt, har teofyllamin intravenöst en säkerställd tilläggs effekt efter behandling med beta-2-stimulerare vid akut svår astma.

Specifik immunterapi

I rapporten framgår att behandling med specifik immunterapi (SIT), som tidigare benämndes hyposensibilisering, har en gynnsam effekt vid allergiutlöst astma, där effekten är väldokumenterad vid allergi mot pollen och katt. Behandling med specifik immunterapi vid allergiutlöst astma och samtidig rinit kommer säkerligen att bli vanligare än vad som är fallet idag. Om denna behandlingsform skall bli mer använd krävs resurser för att ta hand om ett ökat antal patienter.

Hostmedel och antibiotika

Dagligen skrivs många recept på olika hostmedel vid astma och KOL. I SBU-rapporten fastslås att det saknas underlag för att kunna bedöma nyttan av dessa läkemedel. Eftersom många patienter idag undergår långtidsbehandling med acetylcystein bör alltid nyttan av denna utvärderas hos den enskilde patienten.

Idag diskuteras ofta överförskrivning av antibiotika. När det gäller akuta exacerbationer av KOL finns studier som talar för att behandling med antibiotika förkortar dessa.

Av värde för alla allmänläkare

Sammanfattningsvis är SBU-rapporten mycket värdefull som bidrag till att öka kunskapen om behandling av astma och KOL. Den bör ligga till grund för framtida behandlingsrekommendationer från bl a Läkemedelsverket. En ökad följsamhet till SBU:s slutsatser innebär ett bättre utnyttjande av insatta läkemedel sett ur ett hälsoekonomiskt perspektiv samtidigt som patienterna får en effektivare behandling.

Rapporten är givetvis av stort intresse för varje läkare med specialintresse för astma och KOL, men dess slutsatser bör kunna vara till nytta för varje allmänläkare. En rapport som denna blir av nödvändighet mycket omfattande, men den sammanfattning som finns med SBU:s slutsatser är lättillgänglig och lättläst även för en jäktad allmänläkare.

Referens

1. Behandling av astma och KOL. En systematisk kunskapssammanställning. Stockholm: SBU Statens beredning för medicinsk utvärdering 2000; rapport nr 151.

Robert Hahn, professor, överläkare, anestesikliniken, Södersjukhuset

Albumin tvivelaktigt som plasmaexpander inom intensivvård

II Intravenös infusion av albumin framställt ur blodplasma har under 50 års tid använts för behandling av akut hypovolemi. Albumin brukades även länge för att höja plasmans albuminhalt vid hypoalbuminemi. Behandlingens medicinska värde och kostnadseffektivitet har dock ifrågasatts, och indikationerna för albumin har efter hand krympt. Efter en systematisk litteraturöversikt från Cochrane Collaboration kan man till och med fråga sig om albumin alls bör användas. Urvalet av de studier som ingår i analysen kan dock kritiseras.

Infusion av albumin ökar plasmavolymen genom att höja det kolloidosmotiska trycket. Det kroppsegna ursprunget samt att effekten är relativt långvarig, cirka sex timmar, förklarar albuminets framgång på marknaden (Figur 1). Denna plasmafraktion säljs som 3,3-, 5- och 20-procentig lösning i Sverige, varvid den senare ökar blodvolymen med betydligt mer än den givna volymen.

På 1960- och 1970-talen åtnjöt albuminbehandling stor popularitet inom bland annat operations- och intensivvård samt barn- och njurmedicin. Det höga priset på albumin motiverade därefter en viss återhållsamhet och tydligare specificering av preparatens fördelar jämfört med billigare alternativ, såsom kristalloida vätskor och syntetiska kolloider av typ dextran. Fördelar var dock svåra att påvisa. Inom intensivvården enades man om att hypoalbuminemi är ett naturligt fenomen vid svår sjukdom, och överlevnaden tycks inte förbättras av att albumin tillförs [1, 2]. Hypovolemi har därför varit den viktigaste indikationen för albumintillförsel på senare år.

En systematisk översikt som finns tillgänglig på Cochrane Library ifrågasätter även denna indikation för albumin [3]. Analysen fokuserar på överlevnaden hos 1 419 patienter, från 30 studier, som randomiserats till att erhålla eller inte erhålla albumin eller human plasma för behandling av hypovolemi, hypoalbuminemi eller brännskador inom intensivvården. De flesta av studierna är ganska små och hade för avsikt att studera helt andra frågor än överlevnad. Resultatet visar ändå att prognosen försämras när albumin givits. Den relativa risken för död var 1,46 när albumin givits för att behandla hypovolemi, 1,69 om indikationen var hypoalbuminemi och 2,40 om albumin givits vid brännskador. Samtliga delresultat är statistiskt signifikanta eller näst intill signifikanta.

Alla väsentliga delar av denna metaanalys publicerades i British Medical Journal under sommaren 1998 [4]. Där skriver författarna målande att vart 17:e dödsfall i albumingruppen kan förklaras av denna behandling eftersom dödligheten

SAMMANFATTAT

Albumin från blodplasma har länge använts på skilda indikationer inom intensivvården.

Albumin är dock dyrt och därför har användningen minskat.

En översikt från Cochrane Collaboration har visat att användande av albumin som plasmavolymsexpander ökar dödligheten inom intensivvården.

Översikten har följts av en intensiv debatt kring resultatets giltighet samt dess konsekvenser för kliniskt handlande.

Resultatet har tyvärr påverkats av att albumin i flera artiklar givits på ett felaktigt sätt.

Evidensbaserad medicin

var 6 procent högre än hos kontrollerna. Artikeln fick stor uppmärksamhet och dess implikationer diskuterades flitigt på tidskriftens insändarsidor den hösten. I Stockholm anordnade de Läke-medelssakkunnigas (LÄKSAKs) expertgrupp för vätsketerapi och nutrition ett möte med artikelns huvudförfattare, Ian Roberts, som fick svara på frågor. Debatten var tidvis livlig och visade att det kan vara svårt att genomföra en systematisk översikt som accepteras av läkarkåren.

Problem med översikten

Ett första problem är variationen i jämförelsebehandling. I vissa av de ingående 30 studierna jämförs albumin med syntetiska kolloider, i andra har kristalloider givits och i en tredje grupp har kontrollerna inte fått någon vätska alls. Klinikern får således inget alternativ till albumin i sin behandling. Snarast tycks albumin vara sämre än vad som helst. Att albumin skulle vara sämre än ingen volymterapi alls vid hypovolem chock är dock svårt att acceptera för dem som arbetar med sådana patienter.

Ett annat problem är variationen i terapimål. Många av be-



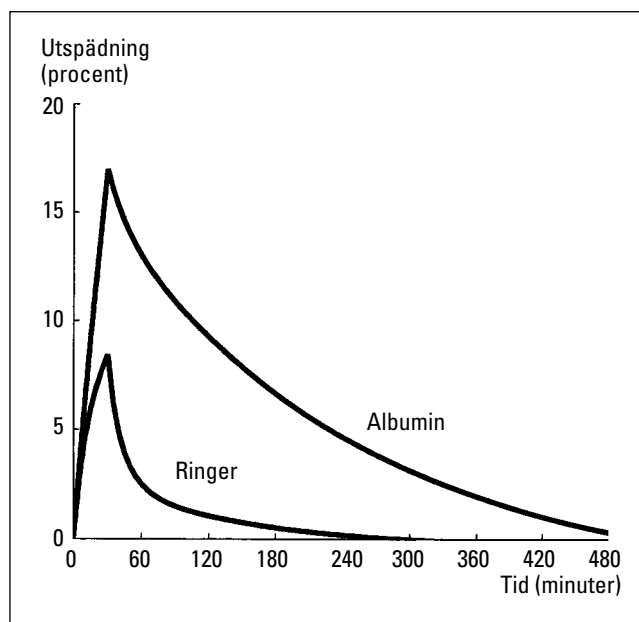
Albumin, 20 procent, ökar blodvolymen med betydligt mer än sin egen volym genom att det kolloidosmotiska trycket är högt. Bilden ställd till förfogande av Annika Hedin, Södersjukhuset.

handlingarna, som kan ha genomförts för mer än 15 år sedan, använde terapimål som idag ter sig olämpliga eller potentiellt skadliga. Ett exempel är att albumin i en studie gavs i mängden 37,5 gram dagligen tills albuminkoncentrationen i plasma normaliserades [5]. Att på detta sätt varje dag öka plasmavolymen med cirka 1 liter hos intensivvårdspatienter torde ha medfört en ansevärd både cirkulatorisk och metabol belastning. Eftersom graden av hypoalbuminemi indikerar en sjukdoms svårighetsgrad torde de svårast sjuka ha belastats mest. Därför är det knappast någon överraskning att dödligheten bara var knappt hälften så hög hos de patienter som slapp denna behandling. Man kan dock fråga sig om resultatet tyder på att albumin är olämpligt eller farligt.

En särskilt stor vikt i sammanställningen av hypovolemistudier har givits åt ett arbete av Lucas och medarbetare som gav 52 traumapatienter »standard of care» utan albumin eller också 150 g albumin under den första operationen samt därefter 150 g albumin per dag under de följande 5 dagarna [6]. Denna behandling ger således patienterna en daglig dos albumin som väl överstiger hela plasmans innehåll av albumin. Denna mängd är knappast tillrädligt att ge i randomiserade studier utan på sin höjd som engångsbehandling vid konstaterad svår hypovolem chock. Ingen torde förvånas över att patienterna dog i albumingruppen.

För flera andra studier kan man också fråga sig om randomiseringen till två i förväg fastställda behandlingar förfuskas möjligheten för behandlande läkare att optimalt behandla patienten. En god hypovolemibehandling kräver fingertoppskänsla och till exempel en återkommande kontroll av patientens hemodynamiska svar för att terapin skall kunna avpassas på ett lämpligt sätt. När man randomiserar patienter till en stor dos albumin givet dagligen kan man som anestesilog uppleva att behandlingen förgrovas jämfört med vad som kan anses vara god klinisk praxis. Detta borde man ta hänsyn till när de enskilda studiernas viktas mot varandra. Den osofistikerade läsaren varnas heller inte vad gäller sådana problem, vilket troligen kan förklaras av att ingen av författarna till denna översikt tycks arbeta inom någon medicinsk specialitet där albumin ofta ges.

Även om den aktuella Cochranestudien kan kritiseras så saknar den konkurrens. Det finns nämligen ingen enskild studie som är tillräckligt stor för att kunna upplysa oss om ett eventuellt samband mellan albuminanvändning och överlevnad efter svår sjukdom, vilket är pinsamt för en behandling som använts mycket och kostat ännu mer. Intensivvård är



Figur 1. Utspädningen av venös plasma under och efter infusion av 10 ml/kg av 5-procentig albuminlösning respektive Ringeracetat på 30 minuter. De datorberäknade kurvorna grundas på medelvärden från volymkinetiska analyser av 15 försökspersoner som erhållit albumin samt 8 som erhållit Ringeracetat [10]. Figuren ställd till förfogande av Annika Hedin, Södersjukhuset.

dock en så komplex verksamhet att få har trots det vara möjligt att påvisa några skillnader i överlevnad beroende på valet av plasmavolymsexpander.

Mekanismen bakom den ökade dödligheten vid albumin-användning är oklar. Ökat hjärtarbete orsakat av hypervolemi samt läckage av albumin från blodbanan till perifera vävnader är troligen bidragande orsaker. Man kan dock undra om tillförseln av kolloiden i sig själv eller om effekter kopplade till albuminmolekylens struktur bär ansvaret. För att belysa i detta publicerade gruppen kring Cochranestudien en annan rapport i mars samma år där man strukturerat nästan samma studier som i analysen av albumin kring riskerna med att använda en kolloid (albumin, plasma eller syntetisk kolloid) med kristalloider av typ Ringeracetat [7]. Resultatet visade att dödligheten var 20 procent hos dem som erhållit kristalloider och 24 procent hos de kolloidbehandlade patienterna. Denna högre dödlighet gällde användning av kolloid vid trauma och brännskador, medan mortaliteten snarast var sänkt vid kirurgi. En metaanalys av en annan grupp, baserad på 17 studier, publicerades några månader senare och fann endast en försämrad överlevnad vid kolloidbehandling av traumapatienter [8].

Albumin i stor dos tvivelaktigt mot hypovolemi

Tidigare författare har visserligen hävdats att albuminbehandling inte erbjuder några specifika fördelar jämfört med andra kolloider [9]. Cochranestudien stödjer detta men gör tillägget att albumin kanske är skadligt. En slutsats man därför kan dra är att rutinmässig tillförsel av albumin för att höja albuminnivån i plasma är olämplig. Det är också tvivelaktigt att ge albumin i stor dos för att motverka hypovolemi hos intensivvårdspatienter. Antalet inkluderade pediatrika studier är dock för begränsat i Cochranestudien för att återhållsamheten med albumin skall utsträckas till denna grupp. Albuminbehandling av hypovolemi som uppstår vid planerade kirurgiska ingrepp kan heller inte bannlysas av andra än ekonomiska skäl.

Till sist har Cochranestudien om albumin tydliggjort att

svagheter kan förekomma hos systematiska översikter trots att de genomförts i enlighet med alla tekniska regler. De studier som väljs ut för att ingå i översikten måste jämföra behandlingar som en god doktor idag skulle kunna tänkas ge sina patienter. I annat fall sjunker översiktens värde snabbt och kan i värsta fall bli helt missvisande som riktlinje för kliniskt handlande.

Referenser

1. Guthrie RD, Hines C. Use of intravenous albumin in the critically ill patient. *Am J Gastroenterol* 1991; 86: 255-63.
2. Marik PE. The treatment of hypoalbuminemia in the critically ill patient. *Heart Lung* 1993; 22: 166-70.
3. The albumin reviewers (Alderson P, Bunn F, Lefebvre C, Li Wan Po A, Lil L, Roberts I, Schierhout G). Human albumin solution for resuscitation and volume expansion in critically ill patient (Cochrane Review). In: *Cochrane Library*, 2000. Issue 1. Oxford: update software.
4. Cochrane injuries group albumin reviewers. Human albumin administration in critically ill patients: systematic review of randomized trials. *BMJ* 1998; 317: 235-40.
5. Golub R, Sorrento JJ, Cantu R, Nierman DM, Moideen A, Stein HD. Efficacy of albumin supplementation in the surgical intensive care unit: a prospective, randomized study. *Crit Care Med* 1994; 22: 613-9.
6. Lucas CE, Weaver D, Higgins RF, Ledgerwood AM, Johnson SD, Bouwman DL. Effects of albumin versus non-albumin resuscitation on plasma volume and renal excretory function. *J Trauma* 1978; 18: 565-70.
7. Schierhout G, Roberts I. Fluid resuscitation with colloid or crystalloid solutions in critically ill patients: a systematic review of randomised trials. *BMJ* 1998; 316: 961-4.
8. Choi PT, Yip G, Quinonez LG, Cook DJ. Crystalloids vs colloids in fluid resuscitation: a systematic review. *Crit Care Med* 1999; 27: 200-10.
9. Zetterström H. Albumin ges fortfarande på tvivelaktiga indikationer. *Läkartidningen* 1997; 94: 711-3.
10. Svensén C, Drobin D, Edsberg L, Ståhle L, Hahn RG. Volymkinetik – en ny metod för att styra intravenös väsketillförsel. *Läkartidningen* 1999; 96: 1969-74.

Särtryck

Läkartidningen

Alla kroppens celler reagerar på olika signalämnen i omgivningen, ämnen som styr deras fundamentala livsprocesser.

Dessa ämnen kallas kollektivt tillväxtfaktorer. En serie i *Läkartidningen* 1995 om dem speglar tendenser i dagens medicinska forskning och pekar på några tillämpningsområden.

Området är i början av en snabb utveckling och många produkter är under utprövning för klinisk användning.

Häftet omfattar 12 artiklar på sammanlagt 56 sidor + färgomslag.

Priset är 90 kronor.



Tillväxtfaktorer

Beställer härmed.....ex
av "Tillväxtfaktorer"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till **Läkartidningen**
Box 5603
114 86 Stockholm

Faxnummer: **08-20 74 35**

www.lakartidningen.se
under särtryck, böcker

Ulf Haglund, professor, institutionen för kirurgi, Akademiska sjukhuset, Uppsala

Evidens om hallux valgus

Klent vetenskapligt underlag för behandling av en så vanlig åkomma

II Evidensbaserade översikter finns av tre principiellt olika typer: systematiska översikter från The Cochrane Collaboration, systematiska översikter publicerade i »peer reviewed«-tidskrifter och systematiska utvärderingar av medicinska teknologier (typ SBUs utvärderingar). Cochranesamarbetet har nyligen givit ut en systematisk översikt avseende behandlingen av hallux valgus [1]. I det följande ges en kommentar till denna.

Målsättningen med systematiska översikter är främst att definiera i vilken utsträckning olika behandlingsformer har stöd i vetenskapligt underlag. Avsaknad av sådant stöd kan bero på att vetenskapliga studier visar att en behandlingsform inte är effektiv, men det kan också bero på att relevanta vetenskapliga studier saknas.

Hallux valgus ett vanligt men besvärande tillstånd

Hallux valgus är ett vanligt förekommande tillstånd som av många patienter upplevs som besvärande, generande och invalidiserande. Åkomman måste dock betraktas som godartad och den tillhör inte de prioriterade sjukdomarna i vårt sjukvårdssystem. Av det skälet blir ofta patienter med hallux valgus bortprioriterade i den offentliga sjukvården och hamnar väntande i långa köer.

Å andra sidan lämpar sig kirurgisk behandling av hallux valgus för dagkirurgi; hallux valgus torde av ovan angivna skäl vara en åkomma som i ganska stor utsträckning omhändertas i privat regi.

Det finns mer än 100 olika beskrivna operationsmetoder för hallux valgus. Det måste därför sägas vara en angelägen uppgift att definiera det vetenskapliga underlaget för en värdering av effektiviteten av olika behandlingsformer. Med tanke på att patientmaterialet är stort skulle man kunna förvänta sig ett relativt gott vetenskapligt underlag för en sådan utvärdering.

En god systematisk översikt behandlar en definierad fråga. Värdering sker på basen av ett noga angivet underlag. Praktiskt innebär det att undersökarna söker efter artiklar som kan tänkas besvara den aktuella frågan i olika i översikten angivna elektroniska databaser. Dessutom kompletteras den elektroniska sökningen oftast med manuell sökning i tidskrifter och böcker där man kan förvänta sig finna artiklar av intresse inom ämnesområdet. Det skall anges i översikten vilka är sökningarna täcker.

Cochrane-samarbetet värderar endast randomiserade kontrollerade studier, men försöker på olika sätt inklude-

SAMMANFATTAT

The Cochrane Collaboration har genomfört en kunskapsöversikt över behandling av hallux valgus. Trots att detta är en vanlig åkomma påträffades bara tolv studier där effekten av behandlingen studerades. Två studier har värderat konservativ behandling (nattlig skena respektive ortos) mot ingen behandling. Det förelåg ingen skillnad mellan dessa alternativ.

I åtta arbeten har olika versioner av kirurgisk behandling jämförts. Debasering (Kellers operation) jämfördes med artrodes respektive osteotomi i studier av låg kvalitet, utan påvisbar skillnad. Chevronosteotomi jämfördes med annan osteotomi, och kirurgens modifikation av standardingrepp jämfördes med standardingreppet i flera studier, utan påvisad skillnad.

Olika postoperativa regimer har jämförts i två studier utan påvisbar skillnad.

Det vetenskapliga underlaget för behandling av denna vanliga åkomma visade sig således kvantitativt och kvalitativt vara klent.

Serie: Evidensbaserad medicin

ra studier publicerade på olika språk. Det klarar man genom att i studiegrupperna inkludera representanter för olika språk.

Cochrane-samarbetet inkluderar även icke publicerade studier. Sådana studier är ju än svårare att spåra men viktiga att ha med för att balansera det faktum att negativa studier tenderar att publiceras mindre ofta. Man kan spåra studier genom att t ex screena ansökningar dels till organisationer som beviljar forskningsanslag, dels till forskningsetiska kommittéer.

Tolv studier omfattande 740 patienter

Den aktuella systematiska översikten om behandling av hallux valgus är utförd av tre briter. De har sökt i stora medi-



cinska databaser, bl a Medline och Embase, och vidare i The Cochrane Controlled Trials Register och The Cochrane Musculoskeletal Injuries Trials Register.

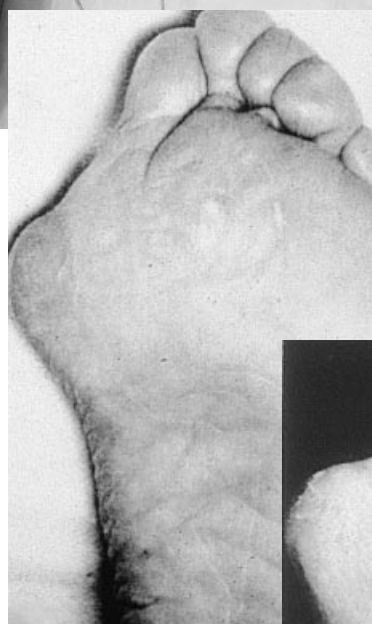
Bibliografier och speciella podiatriska tidskrifter har sökts i manuellt. Registren har sökts från 1966 (Medline), och från 1980 (Embase). Slutdatum för sökningen var oktober 1998.

Underlaget tillfredsställer kriterierna

Det aktuella underlaget tillfredsställer således kriterierna dels att ha en bred sökbaser, dels att det i översikten identifieras hur sökningen gått till. Totalt fångades 20 studier, varav 8 exkluderades.

Orsaken för exkludering finns angiven för varje enskilt fall i den systematiska översikten. Att på basen av titlar och abstrakt fånga studier som kan vara av intresse, och efter genomgång baserad på i förväg uppgjorda kriterier av de så identifierade artiklarna inkludera eller exkludera undersökningar, är metodologiskt kännetecknande för en god systematisk översikt.

Skäl som anges för rejektion i den aktuella studien om hallux valgus är att undersökningen i själva verket visat sig inte vara en randomiserad kontrollerad undersökning, att en preliminär rapport aldrig följts upp eller att studien behandlar andra frågeställningar än behandling av hallux valgus. Tre av de exkluderade studierna behandlade t ex olika anestesiformer



Hallux valgus med uttalad pseudoexostos betingad av prominent caput metatarsale I och bursabildning.

På röntgenbilden ett fall före (t v) respektive efter operation med chevronosteotomi och fixation med skruv.

vid kirurgi för hallux valgus. De 12 efter genomgång inkluderade studierna omfattade totalt 740 patienter.

Varje del poängsätts

Ett viktigt moment i en systematisk översikt är att värdera den vetenskapliga kvaliteten i inkluderade studier. Speciellt tas i den värderingen hänsyn till metoden för randomisering, om behandlingsmetoden varit ordentligt dold för utvärderaren, för behandlande personal och för patienten, om patienter som lämnat studien och skälen härför beskrivits och om analyser gjorts i enlighet med »intention to treat«-metoden.

Definierade inklusions- och exklusionskriterier i studier, definierade interventioner, adekvat tid för uppföljning och i förväg definierade mätningar/värderingar av resultatet från interventionerna är andra områden som inkluderas i utvärderingen av metodkvalitet.

De i denna systematiska översikt ingående studierna har övervägande relativt låg metodkvalitet enligt den utvärdering som gjorts och som rapporteras baserad på ett detaljerat poängssystem omfattande tolv komponenter. Varje del värderas från 0 till 2; som bäst kan således en studie få 24 poäng. I medeltal fick de inkluderade studierna 9 poäng (variationsvidd 4–12), vilket måste anses tyda på en måttlig vetenskaplig kvalitet hos ingående studier.

Objektiva och subjektiva variabler

Det är viktigt att en systematisk översikt fokuserar på variabler som beskriver relevanta resultat då olika studier skall jämföras. I denna översikt har hänsyn tagits till såväl objektiva som subjektiva variabler. De objektiva variablerna har varit hallux valgus-/abduktusvinkeln och första metatarsalvinkeln. I bägge fallen utgörs resultatet av den vinkel som uppmätts röntgenologiskt i slutet av studien. Den har angetts i absoluta tal, inte som procentuell förbättring. Andra objektiva variabler som tagits med i översikten är rörelseomfång i metatarsofalangealleden (omfånget från maximal dorsalflektion till maximal plantarflektion) samt komplikationer. Det sistnämnda mättes i översikten som antalet komplikationer, t ex infektion, behov av reoperation, utebliven läkning och avaskulär nekros, men utan hänsyn till deras svårighetsgrad.

De subjektiva resultatvariabler som inkluderats i denna systematiska översikt inbegriper antalet patienter som vid studiens slut fortfarande hade smärta, graden av patienttillfredsställelse vid samma tidpunkt, förekomsten av kvarvarande begränsningar förknippade med att välja passande fotbeklädnad och förekomsten av inskränkt rörelseförmåga (gångproblem). Alla i översikten ingående studier har inte rapporterat samtliga ovan listade variabler. Därför har jämförelser mellan studiearmar fått baseras på de av ovanstående variabler som medtagits.

Två av studierna beskriver konservativ behandling jämfört med ingen behandling. Totalt inkluderades 150 patienter i de två studierna. Den vetenskapliga metodologiska kvaliteten fick 10 respektive 11 poäng. Med konservativ behandling avses nattlig skena respektive ortos. Denna behandling hade ingen påvisad effekt jämfört med ingen behandling.

I åtta studier utvärderades kirurgisk behandling. I två undersökningar jämfördes debasering (Kellers operation) med andra operationer (artrodes respektive distal osteotomi). Totalt inkluderades 133 patienter i de studierna. Den vetenskapliga kvaliteten värderades lågt (4 respektive 6 poäng enligt ovan). Någon skillnad i resultat mellan operationsmetoderna kunde inte påvisas.

I tre studier, som inkluderade 205 patienter, jämfördes

chevronosteotomi med andra typer av osteotomioperationer. Den vetenskapliga kvaliteten varierade: 7, 10 respektive 12 poäng. Ej heller i den jämförelsen kunde någon fördel med någon av de studerade teknikerna påvisas.

I tre studier, inkluderande 157 patienter, jämfördes standardoperation med kirurgens egen modifikation av metoden. Den vetenskapliga kvaliteten varierade mellan 9 och 11 poäng. I dessa tre studier jämfördes debasering enligt Keller med och utan sentransferering (två studier) respektive Keller med och utan distraktion av leden. Ingen av de tre adaptationerna visade sig innebära någon skillnad för patienten jämfört med standardoperationen. Slutligen jämfördes i två studier olika postoperativa regimer på totalt 95 patienter. Den vetenskapliga kvaliteten värderades till 6 respektive 11 poäng. Tillägg av passiv rörelse efter kirurgi i tidig respektive sen belastning efter kirurgi studerades. Ingen skillnad påvisades.

Bristfälligt vetenskapligt underlag

Det kan tyckas vara lite trist att det vetenskapliga underlaget för behandling av en så vanlig åkomma som hallux valgus kvantitativt är så litet och av så relativt måttlig kvalitet som denna systematiska översikt visar. Författarna till översikten drar slutsatsen att det inte finns tillräcklig evidens för att tillåta evidensbaserad identifiering av bästa behandling av hallux valgus – med undantag för att det inte finns något vetenskapligt stöd alls för konservativ terapi. Det vetenskapliga underlaget för val av operationsteknik är bristfälligt.

Undersökarna konstaterar också att det inte är säkert att röntgenologisk uppföljning med bestämning av hallux valgus-vinkel respektive första metatarsalvinkeln är lämpligaste resultatbeskrivning. Validerade, patientbaserade resultatvariabler bör identifieras och inkluderas i kommande studier. Uppföljningstiden bör, enligt författarna, vara minst tre år.

Pekar på svagheter

Sammanfattningsvis är den av Cochrane-samarbetet stödda systematiska översikten av behandling vid hallux valgus tekniskt välgjord och av hög kvalitet. Den systematiska översikten pekar på svagheter vad gäller klinisk forskning vid en vanligt förekommande kirurgisk åkomma och identifierar klara förbättringsmöjligheter inom området. Denna studie gäller ju endast det vetenskapliga underlaget för behandling av hallux valgus; i strikt mening kan några slutsatser inte dras avseende andra kirurgiska »vardagssjukdomar«.

Emellertid måste man misstänka att slutsatserna är överförbara till alltför många vanliga kirurgiska sjukdomstillstånd. Endast ytterligare systematiska översikter kan ge besked.

Referens

1. Ferrari J, Higgins JPT, Williams RL. Interventions for treating hallux valgus (abductovalgus) and bunions. Oxford: The Cochrane Library, 1999; issue 4: 1-22.

Mats Kalin, docent, infektionskliniken (*mats.kalin@ks.se*)

Magnus Björkholm, professor, hematologiska kliniken; båda Karolinska sjukhuset, Stockholm

Profylaktisk och empirisk behandling mot svamp vid neutropen feber

Referat av metaanalys och kommentar

II Patienter som behandlas med cytostatika för cancer drabbas i stor utsträckning av allvarliga infektioner på grund av såväl försämrat T-cellsförsvar som granulocytopeni. Dessutom skadas slemhinnor i munhåla och gastrointestinalkanal på grund av reducerad nybildning av epitelceller, och barriärfunktionen mot invasion av bakterier, virus och svampar blir därmed defekt. Funktionell neutropeni med ökad infektionsbenägenhet föreligger när neutrofila granulocyter understiger $0,5 \times 10^9/l$. Neutropeni mer än enstaka dagar hos cytostatikabehandlade patienter medför stor risk för invasion av bakterier från munhåla eller tarmkanal med ett ofta fulminant förlopp. Empirisk behandling med bredspektrumantibiotika när en neutropen cancerpatient får feber visades för 20–30 år sedan kunna reducera dödligheten avsevärt och är i dag en hörnsten i behandlingen [1, 2]. Dödligheten i bakteriell infektion vid neutropeni är idag 1–5 procent [1].

Successivt förbättrad cancerbehandling och tillgång till fler och mer potenta antibiotika har medfört att antalet överlevande allvarligt immunosupprimerade cancerpatienter ökar. Oftast behöver patienterna upprepade cytostatikabehandlingar och annan cytoreduktiv terapi för kontroll av sin cancer. Patienter med hematologiska maligniteter, särskilt de som genomgår allogen benmärgstransplantation (BMT), drabbas av svår och långvarig immunosuppression. Risken för invasiva svampinfektioner med ett ej sällan dödligt förlopp är hög hos dessa patienter. Dessutom är diagnostiken bristfällig, och många svampinfektioner diagnostiseras inte förrän efter döden. Svampinfektioner orsakar också betydande morbiditet, till exempel smärta och nutritionssvårigheter vid oral jästsvampsinfektion. Infektioner orsakade av olika typer av svampar utgör därför idag ett allvarligt hot mot framgångsrik cytostatikaterapi vid cancer. Möjligheterna till profylax och tidig empirisk behandling vid misstänkt svampinfektion har därför varit föremål för ett flertal studier och även ett flertal översiktsartiklar [3–7].

Metaanalysens material och metoder

Gøtzsche och Johansen har genomfört en metaanalys av effekten av profylaktisk och empirisk svampbehandling hos neutropena cancerpatienter. Studien publicerades primärt i British Medical Journal 1997 [8], härefter i The Cochrane Li-

SAMMANFATTAT

Gøtzsche och Johansen har för Cochranebibliotekets räkning genomfört en metaanalys av profylaktisk och empirisk effekt av svampbehandling vid neutropen feber omfattande 29 studier med totalt 3 875 patienter.

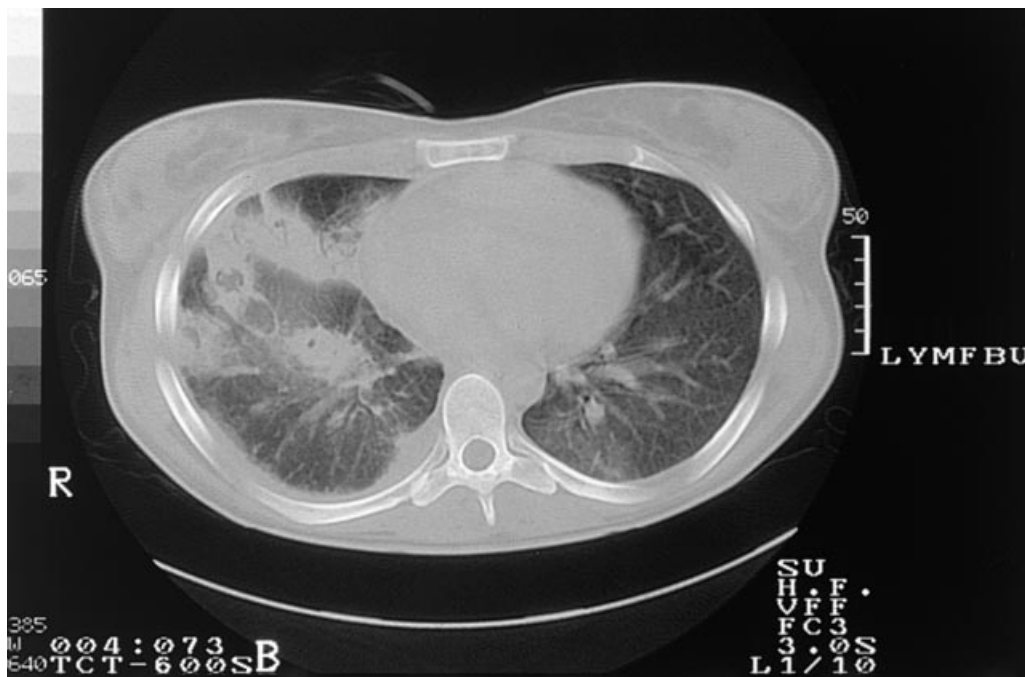
Intravenöst amfotericin B var det enda läkemedel som kunde visas reducera den totala risken för att avlida, och författarna anser att det är detta preparat som bör föredras för profylax och empirisk behandling. I en annan metaanalys har samma författare dessutom funnit att liposomalt amfotericin B är signifikant bättre än konventionellt amfotericin B.

Flukonazol eller ketokonazol minskade påtagligt risken för kolonisation med svamp, och med flukonazol eller itraconazol sågs en halvering av risken för utveckling av invasiv svampinfektion. Någon effekt på risken att avlida kunde inte dokumenteras med dessa preparat.

Vi håller med om att denna metaanalys och andra data ger stöd för att använda någon preparation av amfotericin B vid neutropen feber som inte svarat på bredspektrumantibiotika. Frågan vid vilken tidpunkt sådan terapi ska startas är dock obesvarad. Flukonazolprofylax har sannolikt ingen effekt på den totala risken för död men minskar risken för symtomgivande oral mukositis. Frikostig användning vid symtomgivande mukositis kan vara ett alternativ till profylax.

Evidensbaserad medicin

brary [9] med senaste uppdatering 31 augusti 2000 [10]. Anledningen till uppdateringarna har varit inklusion av ytterligare studier. Detta har medfört att författarna reviderat de ur-



Figur 1. Invasiv aspergillos är en livshotande komplikation hos en patient som behandlats för cancer med kemoterapi eller annan cytoreduktiv behandling.

sprungliga slutsatserna så sent som i augusti 2000 och att presentationen av resultaten lagts upp helt annorlunda med betoning på skillnaderna mellan de mycket olika preparaten. I diskussionen har man tagit in ytterligare en Cochrane-rapport [11], som man anser stöder de nya slutsatserna. Dessutom används relativ risk (RR) i stället för oddskvot (OR) i den senaste rapporten.

Den primära målsättningen har oförändrat varit att undersöka vanligen använda svampmedels effekt på mortalitet. Man har även utvärderat effekten på risken för invasiv svampinfektion och kolonisation. Studier som rörde endast oral candidiasis har dock uteslutits. Denna sjukdom har i stället blivit föremål för en egen Cochrane-rapport [12].

Författarna sökte identifiera alla kontrollerade studier där ett preparat jämförts med placebo eller med ingen behandling. Studier där olika preparat jämförts är alltså inte inkluderade. Först gjordes en sökning i Medline från 1966 och framåt med användning av alla för ämnet aktuella sökord och kombinationer. I de identifierade artiklarna gick man sedan igenom referenslistorna och tog fram de artiklar som inte redan identifierats. Vidare eftersöktes studier genom kontakt med läkemedelsföretag, genom kontakt med författare till publicerade studier som nämnts i publicerade artiklar och genom granskning av abstraktböcker från några utvalda kongresser. Alla författare kontaktades för att bekräfta eller korrigera data, vilket medförde att de slutgiltiga tabellerna i 6 fall skilde sig något från dem som primärt publicerades i de olika studierna. Ett antal kontaktade företag och författare avstod dock från att medverka.

Invasiv svampinfektion definierades som positiv blododling, esofageal Candida-infektion, lunginfektion eller mikroskopiskt verifierad djup svampinfektion (deep tissue infection). Oral eller vulvovaginal svampinfektion, hudinfektioner eller kolonisation av urinen betraktades inte som tecken till invasiv sjukdom. De definitioner av svampinfektion som tillämpades i de olika studierna prövades parallellt för att undersöka hållbarheten i den använda definitionen på invasiv svampinfektion. Effekt på »död i svampinfektion« utvärderades inte eftersom man ansåg att det ofta är svårt att med säkerhet fastställa vilken betydelse en svampinfektion haft för dödsfall. För den statistiska analysen viktades studierna, och

man tog hänsyn till om heterogenitet förelåg mellan de olika studierna. OR med 95-procentigt konfidensintervall användes genomgående i de första versionerna, men i den senaste har man använt RR med konfidensintervall i resultatpresentationen, medan OR fortfarande används i figurerna. Resultaten blir marginellt olika.

Metaanalysens resultat

Man identifierade primärt 31 studier, och härfter har tillkommit ytterligare 5. Studierna granskades av vardera författaren för sig. Olika bedömningar dem emellan löstes diskussionsvägen. Sju studier exkluderades på grund av bristfällig randomisering, avsaknad av neutropeni hos merparten patienter eller ofullständiga data. Vid senaste uppdateringen återstod 29 studier, inkluderande 3 875 patienter. Leukemi var dominerande underliggande diagnos i 18 studier och benmärgstransplantation i 11. Tjugosex studier behandlade profylax mot svampinfektioner och 3 empirisk behandling av misstänkt svampinfektion. I den initiala rapporten uppgavs en medianuppföljningstid av 56 dagar i de 13 studier där uppföljningstid var angiven, men någon redovisning av detta finns inte i den senaste uppdateringen.

Den totala mortaliteten var 15,6 procent (544/3 496), baserat på de studier där mortalitet var angiven. Enligt den ursprungliga rapporten förelåg det ingen heterogenitet mellan studierna och ingen skillnad i effekt mellan leukemi- och BMT-patienter. Man kunde inte notera någon signifikant effekt på total mortalitet av profylaktisk eller empirisk svampbehandling om alla preparaten behandlades tillsammans (RR 0,94, CI 0,81–1,10). Det förelåg emellertid klara skillnader mellan de olika preparaten, vilka således lyfts fram i den senaste versionen.

Med amfotericin B-profylax eller empirisk behandling var mortaliteten 47/425 (11,1 procent) medan den i kontrollgruppen var 63/431 (14,6 procent). Den relativa risken för död var 0,72 (CI 0,51–1,02, P=0,06). I de inkluderade åtta studierna hade patienterna behandlats med peroral profylax (1 studie), med lågdos (0,1 mg/kg) intravenös profylax (2), med liposomalt amfotericin B för profylax (2) och empiriskt för misstänkt svampinfektion (3). I de tre studierna i vilka effekten av empirisk behandling med amfotericin B undersöktes, var den

Tabell I. Effekt av olika antimykotiska medel jämfört med placebo på fyra olika effektvariabler. Oddsquot (OR) och 95-procentigt konfidensintervall anges (inom hakparentes anges även antal studier/antal patienter).

Preparat	Effekt på risk för död	Effekt på risk för invasiv sjukdom	Effekt på risk för kolonisation	Effekt på risk för »rescue therapy«
Amfotericin B	0,67 (0,44–1,01) [8/856]	0,37 (0,19–0,71) [7/723]	0,42 (0,26–0,70) [3/378]	0,53 (0,34–0,84) [5/504]
Flukonazol	1,05 (0,80–1,39) [7/1 470]	0,38 (0,26–0,55) [8/1 539]	0,34 (0,26–0,43) [6/1 401]	0,73 (0,59–0,90) [7/1 469]
Itrakonazol	0,79 (0,46–1,38) [2/503]	0,43 (0,19–0,97) [2/503]	0,88 (0,49–1,59) [2/503]	0,63 (0,41–0,97) [1/405]
Ketokonazol	0,96 (0,56–1,64) [4/429]	1,37 (0,66–2,83) [7/562]	0,40 (0,28–0,57) [8/626]	0,73 (0,47–1,14) [6/412]
Mikonazol	1,21 (0,65–2,25) [2/238]	0,51 (0,20–1,30) [2/238]	1,44 (0,86–2,41) [2/238]	1,44 (0,84–2,47) [2/238]
Itrakonazol/ketokonazol/amfotericin B			0,31(0,08–1,27) [1/40]	
Totalt	0,93 (0,77–1,12) [23/3 496]	0,47 (0,36–0,61) [26/3 605]	0,45 (0,38–0,54) [22/3 186]	0,73 (0,62–0,85) [21/3 028]

Antal studier och antal patienter är olika för olika effektvariabler eftersom inte alla sådana studerats för alla patienter i alla studier. OR har använts genomgående i tabellen för att få den fullständig. Götzsche och Johansen använder i den senast uppdaterade versionen av rapporten relativ risk (RR), då man anser detta vara mer korrekt. Emellertid finns inte RR för alla preparat och effektvariabler, och skillnaderna mellan OR och RR är små. Fetstil för OR anger signifikant effekt. För amfotericin B:s effekt på risken för död har fetstil använts eftersom Götzsche och Johansen anser att effekten får anses dokumenterad.

sammanlagda OR (finns bara i den ursprungliga rapporten, RR ej beräknad) 0,67 (CI 0,33–1,38), men antalet dödsfall bara 15 vs 20. Beträffande övriga preparat kunde ingen effekt på risken för död noteras (Tabell I).

Enligt den ursprungliga analysen var den totala effekten av givna antimykotika en mycket signifikant ($P < 0,00001$) effekt på risken för invasiv svampinfektion både med metaanalysens definition (OR 0,47, CI 0,36–0,61) och med författarnas egna definitioner (OR 0,43, CI 0,32–0,57). Amfotericin B, flukonazol och itraconazol uppvisade en effekt av jämförbar storleksordning på risk för invasiv svampinfektion och risk att »rescue therapy« skulle behöva användas (Tabell I). Med amfotericin B, flukonazol och ketokonazol sågs en påtaglig reduktion av risken för kolonisation med svamp. Mikonazol saknade effekt avseende alla fyra studerade effektvariabler.

Metanalysens konklusion och diskussion

Författarna diskuterar vilka felkällor och vilken risk för fel-tolkningar som kan föreligga. De framhåller total mortalitet som viktigaste effektvariabel. Författarna har avstått från att registrera död i svampinfektion eftersom man ansett att detta är mer osäkert och mindre relevant än total mortalitet, en uppfattning som vi delar. Död i svampinfektion kan vara associerad med död av andra orsaker. Uppföljningstid och sätt att registrera mortalitet varierade påtagligt mellan de olika studierna. Beträffande en studie föreligger till och med diskrepans mellan data rapporterade i ett konferensabstrakt och i den slutliga rapporten på grund av ändring av uppföljningstiden.

Informella interimanalyser, utan klara regler för hur resultaten av dessa skulle användas, tillämpades i flera studier, vilket kan medföra risk för fel. Interimanalyser rapporteras också bristfälligt. Publiceringsbias har troligen förekommit. En studie med flukonazol stoppades t ex av okänd anledning under studiens gång och inga resultat blev publicerade. Resultat i sammanfattningar till kongresser och i de slutliga rapporterna skilde sig i några fall. Försök till kontakt med författare och företag gav inte alltid klarhet. Speciellt företagets påtagliga ovillighet till samarbete för att få fram så korrekta resultat som möjligt till grund för denna typ av metaanalys är nedslående.

Författarna anser att intravenöst amfotericin B är det enda läkemedel som bevisats kunna reducera den totala risken för

att avlida och att det därför är det preparat som bör föredras för profylax och empirisk behandling. Beträffande övriga preparat anser man inte att det föreligger tillräckligt med data för att bedöma deras värde.

I diskussionen för författarna i resultaten från tre stora, nyligen publicerade studier i vilka man jämfört liposomalt amfotericin B (AmBisome) med konventionellt. Denna analys har blivit föremål för en egen Cochranerapport [11]. Totalt förekom 118 dödsfall i dessa tre studier, och den relativa risken för död var efter behandling med liposomalt amfotericin B 0,74 (CI 0,52–1,07) jämfört med konventionellt amfotericin B. Författarna har kombinerat resultaten från denna metaanalys [11] med den här refererade analysen, där ju konventionellt amfotericin B jämförts med placebo eller med ingen terapi [10]. Man erhåller på detta sätt en relativ risk av 0,71 (CI 0,56–0,91) och beräknar därifrån att man skulle behöva behandla 25 patienter med intravenöst amfotericin B för att rädda ett liv.

Någon effekt av flukonazolprofylax på den totala mortaliteten kunde inte påvisas trots att denna del av metaanalysen är den största med totalt 1 539 inkluderade patienter. I alla studierna användes hög dos, 400 mg per dag. Rescue therapy med amfotericin B vid misstanke om svampinfektion användes totalt sett oftare i de obehandlade grupperna (OR 0,73, CI 0,62–0,85). Beslut att inleda behandling med ett nytt svampmedel kan naturligtvis påverkas av förekomst av oral candidiasis eller fynd av svamp i odlingar, vilka enligt studiens resultat var vanligare hos kontroller. För tidigt startad rescue therapy i kontrollgruppen skulle kunna medföra att en effekt av flukonazol på mortalitet förbisågs. Detta anser författarna inte vara troligt och anför en registrerad effekt på mortalitet i de studier där amfotericin B jämförts med placebo trots att rescue therapy startades lika tidigt i kontrollgrupperna i dessa studier som i flukonazolstudiernas kontrollgrupper.

Författarna diskuterar möjligheten av att det skulle föreligga andra effekter av imidazolpreparat. Man framhåller att ketokonazol befunnits vara immunsupprimerande och förenat med ökad frekvens bakteriemi i tre olika studier (totalt 70 vs 45 jämfört med placebo i två studier [13, 14], 20 vs 15 procent i den tredje [15]). Möjligen skulle detta kunna vara en klasseffekt för denna grupp av svampmedel: beträffande flukonazol förelåg i två studier bakteriemi i 42 vs 23 fall [16, 17]

och beträffande itraconazol i 47 vs 31 [18]. Flukonazol har också associerats med ökad risk för transplanterad kontra värd-sjukdom (GVHD) och organsvikt med 44 vs 23 dödsfall [19] och 102 vs 85 fall av GVHD [20]. I Goodmans studie förelåg ingen skillnad i total mortalitet mellan flukonazol och placebo, men en signifikant skillnad i död i invasiv svampinfektion (1 vs 10, $P=0,01$), vilket enligt Götzsche och Johansen innebär att död av andra orsaker faktiskt var signifikant vanligare bland dem som fått flukonazol än bland dem som fått placebo (54 vs 36, $P=0,04$).

I den här refererade metaanalysen befanns amfotericin B och flukonazol ha jämförbar effekt vad avser reduktion av risken för invasiv svampinfektion. Författarna påpekar att effekten av dessa två preparat jämförts i mer än ett dussin studier, men att det ändå är det svårt att uttala sig om huruvida det föreligger någon skillnad mellan dem. Författarnas försök att sammanställa resultaten i en metaanalys försvårades avsevärt av studiernas uppläggning [21, 22]. I de flesta av studierna har man använt peroralt amfotericin B, som inte absorberas från gastrointestinalkanalen. Vidare har i tre stora studier resultaten för amfotericin B kombinerats med dem för nystatin, som förefaller sakna effekt.

Effekten av olika preparat på kolonisation diskuteras inte.

Kommentar

Osäkra svar på viktiga frågor

I de versioner som har publicerats före den senaste uppdateringen i augusti 2000 [9] ville Götzsche och Johansen avfärda den utbredda uppfattningen att tidigt insatt antimykotisk behandling skulle vara avgörande för prognosen vid neutropen feber som inte svarat på bredspektrumantibiotika. Denna slutsats baserades på resultaten från hela metaanalysen inkluderande alla de olika preparaten. Författarna bortsåg från den redan då iakttagna effekten av amfotericin B och ansåg att svampbehandling skulle användas endast vid diagnostiserad infektion.

I den senaste versionen [10] framför författarna att intravenöst amfotericin B är det enda preparat som visats reducera risken för död i svampinfektion och därför är det preparat som bör användas för profylax och empirisk behandling. Beträffande övriga preparat anser man att det inte föreligger tillräckligt med data. Författarna framhåller att effekt av amfotericin B profylaktiskt eller vid empirisk terapi påvisats i alla publicerade studier utom en. Effekten av amfotericin B är emellertid inte så lätt att värdera. I den här refererade metaanalysen inkluderades en studie med peroralt preparat (som inte absorberas – det är alltså ett helt annat koncept), två profylaxstudier med lågdos amfotericin B (0,1 mg per kg), en profylaxstudie med liposomalt amfotericin B samt tre studier med användning av preparatet för empirisk behandling. I diskussionen för författarna så i resultaten från tre studier i vilka konventionellt amfotericin B jämförts med liposomal beredning (AmBisome) [23-25]. Kombinationen av alla dessa studier används för att beräkna att 25 patienter skulle behöva behandlas med intravenöst amfotericin B för att rädda ett liv. Denna siffra måste betraktas som osäker.

Författarnas slutsats att amfotericin B bör användas för empirisk behandling vid misstänkt svampinfektion är emellertid i linje med den allmänna uppfattningen och den idag använda rutinen på centra där hematologiska maligniteter behandlas. Underlaget för empirisk användning är dock fortfarande lika begränsat som det varit i tio år: endast tre studier

med ett totalt sett begränsat antal dödsfall [26-28]. Någon signifikant effekt på mortaliteten erhöles inte om bara dessa tre studier beaktades. Användning av empirisk svampbehandling har grundats på trender och subgruppsanalyser i framför allt två av dessa studier [29]. Med hänsyn till bristande diagnostisk precision och hög dödlighet vid etablerad svampinfektion är rutinen tidig empirisk behandling dock fortfarande nödvändig.

Tendensen under senare år har varit att sätta in empirisk behandling mot svampinfektion allt tidigare utan att man har några klara hållpunkter för att detta är motiverat. Nya studier har lagts upp så att man inkluderat patienter för empirisk antimykotisk terapi vid neutropen feber som inte svarat på 3-4 dagars behandling med bredspektrumantibiotika. Man har inte tagit hänsyn till om patienterna trots allt är under förbättring. Hur stor andel av dessa patienter som verkligen har nytta av antimykotisk behandling är okänt. Man får dock hålla med Götzsche och Johansen om att de likartade trenderna i studier som jämför amfotericin med placebo och i studier som jämför liposomalt amfotericin med konventionellt talar för att empirisk behandling mot svamp är berättigad. Vi har dock inte kommit närmare svaret på den viktiga frågan när behandlingen ska sättas in, inte heller hur många som verkligen har nytta av behandlingen. Den senare frågan hänger samman med kostnaden eftersom liposomalt amfotericin B är förenat med betydligt färre biverkningar och sannolikt bättre effekt, men för närvarande till mycket höga kostnader.

Imidazolpreparatens användning är inte lättare att ta ställning till. Miconazol saknar effekt, och ketokonazol är belastat med biverkningar och interaktionsproblem [4]. Dessa preparat används därför inte längre för profylax. Flukonazol används frikostigt som profylax mot svampinfektioner vid neutropeni. En viktig anledning är att man vill reducera risken för symtomgivande mukositer med smärtor och nutritionssvårigheter. För detta ändamål har flukonazol befunnits effektivt, vilket stöds i en metaanalys av behandling av mukositer och orofaryngeal candidiasis [12]. Man vet också att höggradig kolonisation med jästsvamp ökar risken för invasiv infektion [30, 31]. Det är dock störande att man inte kunnat påvisa någon effekt på total mortalitet. Därför har Infectious Disease Society of America (IDSA) inte funnit anledning rekommendera rutinmässig svampprofylax vid neutropeni [2]. Man rekommenderar inte heller profylax med annat preparat annat än i speciella situationer, däremot empirisk behandling med något amfotericin B-preparat vid neutropen feber som inte svarar på antibiotikabehandling. Eventuell negativ effekt av imidazoler kan mycket väl vara dosberoende. Data som kan belysa detta saknas dock. Metaanalysens resultat och övriga tillgängliga data talar troligen mot rutinmässig användning av flukonazol för profylax, men för relativt frikostig användning vid symtomgivande mukositer och vid höggradig kolonisation [3, 6, 12]. Huruvida flukonazol kan vara ett alternativ till amfotericin B-preparationer för empirisk behandling av antibiotikaresistent neutropen feber är ännu oklart, men något stöd ges inte i denna metaanalys även om jämförande studier nu finns.

Fördelen med itraconazol är att preparatet har effekt på mögelsvampar, speciellt Aspergillus, vilket inte flukonazol har. På centra där man har problem med aspergillos kan itraconazol därför vara ett alternativ. I ännu högre grad gäller detta patienter som tidigare haft aspergillos och ska få ny cytostatikakur, då man vet att reaktivering är mycket vanlig i dessa situationer.

Såsom förhoppningsvis framgått föreligger betydande skillnader mellan de studier som inkluderats i den här refererade metaanalysen, framför allt i de undersökta antimykotiska preparatens egenskaper, men även i upplägg och målsätt-

ningar i de olika studierna. De uppnådda slutsatserna måste därför betraktas med viss försiktighet. För kommande studier skulle man önska sig likformig rapportering av dödlighet (författarna har använt 90 dagar när så varit möjligt), inte minst eftersom metaanalyser även framgent får förmodas bidra till utvärderingen av effekter med antimykotika. Meningsfulla metaanalyser är ju särskilt angelägna när det gäller sjukdomar med låg incidens, såsom invasiva svampinfektioner.

*

Pfizer Sverige, som marknadsför flukonazolpreparatet Diflucan Pfizer, har erbjudits kommentera artikeln och har svarat att man inte har några synpunkter på författarnas slutsatser eller kommentarer.

Referenser

1. Klastersky J. Science and pragmatism in the treatment and prevention of neutropenic infection. *J Antimicrob Chemother* 1998;41 Suppl D:13-24.
2. Hughes WT, Armstrong D, Bodey GP, Brown AE, Edwards JE, Feld R, et al. 1997 guidelines for the use of antimicrobial agents in neutropenic patients with unexplained fever. *Infectious Diseases Society of America. Clin Infect Dis* 1997;25(3):551-73.
4. Björkholm M. Chemoprophylaxis of fungal infections in neutropenic patients. *Ann Oncol* 1994;5(7):571-4.
5. Lortholary O, Dupont B. Antifungal prophylaxis during neutropenia and immunodeficiency. *Clin Microbiol Rev* 1997;10(3):477-504.
6. Uzun O, Anaissie EJ. Antifungal prophylaxis in patients with hematologic malignancies: a reappraisal [see comments]. *Blood* 1995;86(6):2063-72.
7. Warnock DW. Fungal infections in neutropenia: current problems and chemotherapeutic control. *J Antimicrob Chemother* 1998;41 Suppl D:95-105.
8. Gotzsche PC, Johansen HK. Meta-analysis of prophylactic or empirical antifungal treatment versus placebo or no treatment in patients with cancer complicated by neutropenia [see comments]. *BMJ* 1997;314(7089):1238-44.
10. Gotzsche PC, Johansen HK. Routine versus selective antifungal administration for control of fungal infections in patients with cancer (Cochrane review). *Cochrane Database Syst Rev* 2000;4.
11. Johansen HK, Gotzsche PC. Amphotericin B lipid soluble formulations vs amphotericin B in cancer patients with neutropenia (Cochrane review) [In Process Citation]. *Cochrane Database Syst Rev* 2000;3.
12. Clarkson JE, Worthington HV, Eden OB. Prevention of oral mucositis or oral candidiasis for patients with cancer receiving chemotherapy (excluding head and neck cancer). *Cochrane Database Syst Rev* 2000;2.
19. Goodman JL, Winston DJ, Greenfield RA, Chandrasekar PH, Fox B, Kaizer H, et al. A controlled trial of fluconazole to prevent fungal infections in patients undergoing bone marrow transplantation [see comments]. *N Engl J Med* 1992;326(13):845-51.
20. Slavin MA, Osborne B, Adams R, Levenstein MJ, Schoch HG, Feldman AR, et al. Efficacy and safety of fluconazole prophylaxis for fungal infections after marrow transplantation – a prospective, randomized, double-blind study. *J Infect Dis* 1995;171(6):1545-52.
23. Leenders AC, Daenen S, Jansen RL, Hop WC, Lowenberg B, Wijermans PW, et al. Liposomal amphotericin B compared with amphotericin B deoxycholate in the treatment of documented and suspected neutropenia-associated invasive fungal infections [see comments]. *Br J Haematol* 1998;103(1):205-12.
24. Walsh TJ, Finberg RW, Arndt C, Hiemenz J, Schwartz C, Bodenheimer D, et al. Liposomal amphotericin B for empirical therapy in patients with persistent fever and neutropenia. *National Institute of Allergy and Infectious Diseases Mycoses Study Group* [see comments]. *N Engl J Med* 1999;340(10):764-71.
25. Prentice HG, Hann IM, Herbrecht R, Aoun M, Kvaloy S, Catovsky D, et al. A randomized comparison of liposomal versus conventional amphotericin B for the treatment of pyrexia of unknown origin in neutropenic patients. *Br J Haematol* 1997;98(3):711-8.
26. EORTC. Empiric antifungal therapy in febrile granulocytopenic patients. *EORTC International Antimicrobial Therapy Cooperative Group. Am J Med* 1989;86(6 Pt 1):668-72.
27. Goldstone AH, O'Driscoll A. Early AmBisome in febrile neutropenia in patients with haematological disorders. *Bone Marrow Transplant* 1994;14(Suppl 5):S15-7.
28. Pizzo PA, Robichaud KJ, Gill FA, Witebsky FG. Empiric antibiotic and antifungal therapy for cancer patients with prolonged fever and granulocytopenia. *Am J Med* 1982;72(1):101-11.
30. Martino P, Girmenia C, Venditti M, Micozzi A, Santilli S, Burgio VL, et al. Candida colonization and systemic infection in neutropenic patients. A retrospective study. *Cancer* 1989;64(10):2030-4.
31. Bow EJ, Loewen R, Cheang MS, Schacter B. Invasive fungal disease in adults undergoing remission-induction therapy for acute myeloid leukemia: the pathogenetic role of the antileukemic regimen. *Clin Infect Dis* 1995;21(2):361-9.

I Läkartidningens elektroniska arkiv
<http://lartarkiv.lakartidningen.se>
 är artikeln kompletterad med fullständig referenslista.

SUMMARY

Prophylactic and empirical treatment of fungal infection in neutropenic fever
 Report and comment on a meta-analysis

Mats Kalin, Magnus Björkholm
Läkartidningen 2001;98: 1899-1903

Patients treated for cancer with chemotherapy and other cytoreductive therapy often develop serious bacterial, viral, and fungal infections due to B- and T-cell depletion, neutropenia and decreased barrier function of mucosal membranes. In patients with neutropenic fever not responding to broad spectrum antibiotic therapy there is a high risk of fungal infection. In a Cochrane Library meta-analysis of the effect of prophylactic and empirical antifungal treatment, amphotericin B was found to be the only drug affecting total mortality. Fluconazole, ketoconazole and itraconazole affected risk of colonization and invasive infection, but not the total risk of death. We agree with these conclusions, but emphasize that the appropriate time to initiate empirical antifungal therapy is still not settled. We also believe that azoles may be of value in the therapy of symptomatic mucositis.

Correspondence: Mats Kalin, Dept of Infectious Diseases, Medical Division, Karolinska sjukhuset, SE-171 76 Stockholm, Sweden (mats.kalin@ks.se)

Kisslarmet försvarar sin plats

Kommentar till en metaanalys av larmbehandling vid nattväta

II Med en prevalens på över 6 procent hos sjuåringar är nattväta (enures) ett av våra vanligaste och socialt mest stigmatiserande dolda handikapp. Nattvätans konsekvenser för växande människors mentala hälsa och sociala samspel i stort är självklart svåra att värdera, men det är visat att barn som kissar på sig har ett klart sämre självförtroende än torra barn och att denna skillnad prompt försvinner vid framgångsrik behandling av nattvätan [1]. Att lida av kroniskt dåligt självförtroende under skolåren är knappast ofarligt.

Därför är det besvärande att vi ännu är så dåliga på att behandla nattväta. Vi känner idag inte till någon behandlingsmetod som ger det drabbade barnet mer än kanske 65 procents chans till torra nätter. Det här beror, vill jag hävda, till stor del på att forskningen kring nattvätans patogenes ända fram till de senaste decennierna varit ensidigt fokuserad på psykiatriska förklaringsmodeller. Förutom att detta lett till en ökad skuldbeläggning av barnen och deras föräldrar har det inte givit upphov till några behandlingsmetoder med bevisad effekt.

Senare tids forskning kring nattvätans orsaker har visat att tre faktorer, i olika kombinationer, är centrala: sömnen [2], urinproduktionen [3] och blåsfunktionen [4]. Barn med nattväta kissar på sig för att de är svåra att väcka och för att de dessutom producerar en överdrivet stor mängd urin nattetid och/eller har en liten och hyperaktiv urinblåsa [5].

Enureslarmet utvecklades under senare hälften av 1900-talet som ett alternativ till psykoterapi. Det är fascinerande att själva titeln på den artikel i vilken den första regelrätta larmbehandlingen beskrivs ger uttryck för föreställningen om nattväta som en själslig åkomma [6]. Terapiformen har dock visat sig överleva sin tids synsätt. Idag är det mycket få seriösa barnmedicinska auktoriteter som hävdar själsliga faktorer som huvudorsaker till nattväta, medan larmet fortfarande är ett av två förstahandsalternativ.

Det typiska enureslarmet består av en urindetektor – placerad under lakanen eller i barnets pyjamasbyxor – som är ansluten till en stark ringklocka eller annan anordning som ska väcka barnet. Principen bakom de olika enureslarmen är skenbart enkel: genom att barnet konsekvent väcks i samband med den nattliga blåstömningen lär han eller hon sig gradvis att i stället vakna av sig själv innan tömningen har skett. Nu störs denna enkla förklaring dels av att man kan ifrågasätta om det verkligen går att lära sig något när man sover, dels av att det visat sig att många barn som framgångsrikt behandlats med enureslarmet inte därefter får vanan att vakna och gå på toaletten på nätterna utan i stället sover ostört natten igenom i en torr säng. Under alla omständigheter har larmet vunnit en

SAMMANFATTAT

Enureslarmet har under många decennier varit en förstahandsbehandling vid nattväta (enures), men den vetenskapliga utvärderingen har varit bristfällig. Därför genomfördes nyligen via Cochranesamarbetet en systematisk litteraturöversikt avseende enureslarm.

Sammanlagt fann man 52 studier i vilka enureslarmet ingick som ett behandlingsalternativ. Av dessa exkluderades 30 från den slutgiltiga analysen på grund av metodologiska brister. Resultaten i de fem studier som jämförde enureslarmet med ingen behandling talade entydigt för att larmet var effektivt.

Arbeten som jämförde olika typer av enureslarm med varandra visade inga skillnader. Inget stöd gavs för att larmbehandling skulle bli effektivare om den kombinerades med andra icke-farmakologiska behandlingsmetoder. Det fåtal studier som jämförde larmet med farmakologisk behandling visade inte att det ena var bättre än det andra, åtminstone inte på kort sikt.

Konklusion: enureslarmet förblir en effektiv metod att behandla nattväta, men mer forskning behövs för att visa vilka patientgrupper som är mest gynnade av den ena eller den andra behandlingen för sin nattväta.

Evidensbaserad medicin

framskjuten plats i behandlingsarsenalen vid nattväta. Cochranebiblioteket har nyligen publicerat en systematisk litteraturöversikt och metaanalys avseende larmbehandling vid nattväta hos barn [7]. Syftet med metaanalysen var att bedöma enureslarmets terapeutiska effekter vid nattväta och att jämföra denna behandlingsform med andra interventioner. Syftet med denna artikel är att kommentera metaanalysen och utifrån detta försöka dra slutsatser om enureslarmets plats i behandlingen av nattväta i Sverige.

II Den systematiska litteraturgenomgången

Utförliga sökningar genomfördes i samtliga relevanta databaser, Medline, Embase, PsycLIT m fl, varvid alla potentiellt

användbara artiklar, oavsett språk, genomlästes och referenslistorna i dessa studerades i sökandet efter ytterligare rapporter. Tillverkare av enureslarm och forskare i området kontaktades också. Sökningen gäller artiklar publicerade (eller på annat sätt redovisade) till och med år 1997.

I den slutgiltiga analysen inkluderades endast randomiserade kontrollerade studier i vilka de deltagande barnens sängvätningsfrekvens före behandling (baslinje) dokumenterats och bakomliggande somatiska sjukdomar uteslutits. Behandlingsframgång uttrycktes antingen i reduktion av antalet våta nätter per vecka under eller efter behandlingen, eller i andel barn som uppnått 14 torra nätter i följd.

Studiernas kvalitet bedömdes utförligt i olika avseenden. Således granskades randomiseringsproceduren, kontroll- och behandlingsgruppernas likvärdighet före behandling, redovisningen av bortfallet och av antalet potentiellt inkluderbara barn som inte deltagit, definitionen av effektvariabeln, de statistiska metoderna och uppföljningstiderna. Av inklusionskriterierna för de enskilda studierna uttröntes specifikt huruvida barn med dagtida urininkontinens exkluderats eller inte. Två oberoende bedömare granskade alla artiklar.

Resultat

Sökningen resulterade i att man fann 52 studier i vilka enureslarmet ingick som ett behandlingsalternativ. Av dessa exkluderades 30 studier då de inte var randomiserade kontrollerade studier, inte presenterade baslinjedata eller inte hade uteslutit barn med möjlig bakomliggande somatisk sjukdom. Således ingick 22 studier – eller 1 125 barn – i den slutgiltiga analysen. För ytterligare fyra potentiellt relevanta studier uppges att bedömning ännu pågår. Värt att notera är att det i bara fyra av de 22 studierna uttryckligt specificerats att barn med samtidig dagtida urininkontinens inte ingick [8-11].

Ingen av de analyserade studierna var av helt fläckfri metodologisk kvalitet. Vad som framför allt brast var randomiseringens genomförande och frånvaron av utvärdering enligt »intention-to-treat«-principer, dvs med bortfallet medräknat. De flesta studierna hade dock likvärdiga patient- och kontrollgrupper.

I fem studier (fyra artiklar) jämfördes enureslarmet med ingen behandling [12-15]. Chansen att uppnå 14 torra nätter i rad var fyra gånger större med än utan larm (relativ risk att inte uppnå 14 torra nätter: 0,27; 95 procents konfidensintervall 0,19–0,39).

I fem små studier jämfördes olika typer av larm (larm med detektor i kroppsburen eller placerad under lakanen, larm med direkt signal eller med signal fördröjd några sekunder), men inga signifikanta skillnader kunde noteras.

Tre studier jämförde enureslarmet med andra mer eller mindre beteendeterapeutiska behandlingar, såsom rituell bestraffning(!) [16] och väckningsschema, varvid larmet gav störst chans att uppnå torrhet. Vid nio jämförelser av isolerad larmbehandling och larmbehandling kombinerad med olika beteendeterapeutiska tillägg såsom överinläring, blåstjningsövningar eller nattliga väckningar vid bestämda klockslag noterades inga skillnader vad gäller chansen att bli torr.

Sex av de studier som uppfyllde genomgångens urvalskriterier var jämförelser mellan enureslarmet och olika läkemedel. I fyra av dessa var ett eller flera av jämförelseläkemedlen tricykliska antidepressiva, som ju förut var en vanlig behandling vid nattväta. Patientgrupperna var små, och inga klara skillnader noterades vad gäller behandlingseffekten, däremot föreföll risken för återfall vara mindre efter larmbehandling. Endast i Søren Willes studie från 1986 [11] jämfördes enureslarmet med desmopressin. Här noterades att behandlingssvaret kom snabbare med läkemedlet men att svaret var något bättre med larmet än med desmopressin efter tre månaders be-

handling och att återfallsrisken var påtagligt högre efter desmopressinbehandling.

Tre små jämförande studier av isolerad larmbehandling och larmbehandling kombinerad med läkemedel gav vid handen att både desmopressin- och imipraminbehandling kan tänkas öka larmbehandlingens framgång.

En analys av studier som, på grund av bristande baslinjedata eller osäker exklusion av barn med organisk orsak till sin nattväta, inte uppnått inklusionskriterierna för den egentliga metaanalysen redovisades också. Här tillkom fem studier av enureslarm jämfört med ingen behandling. Samtliga dessa gav resultat i paritet med ovan nämnda studier, alltså en tydligt positiv behandlingseffekt av larmet.

I det fåtal studier som redovisat förekomsten av oönskade effekter av larmbehandling nämndes tekniska problem som falsklarm hos uppemot tre fjärdedelar av deltagarna, medan rädsla förekom hos en liten minoritet av barnen. Fenomenet att barnet inte vaknar av larmet, men att resten av hushållet gör det, inträffade hos mer än hälften av familjerna.

Konklusioner

Författarna till den systematiska litteraturgenomgången drar slutsatsen att behandling med enureslarm är klart effektivare än ingen behandling. Detta är den enda helt säkra slutsats som materialet tillåter. I övrigt ger studierna grund för antagandet att enureslarm fungerar bättre än väckningsschema och att metoden ger mindre återfallsrisk än medikamentell behandling (desmopressin eller tricykliska antidepressiva). Inga säkra svar kan dock ges på frågan om larmbehandling eller läkemedel har störst chans att ge barnet torra nätter på kort sikt. Bedömningen av de flesta studierna störs av att den samtidigt förekomsten av dag- och nattväta inte redovisats. Barn med associerad dagtida urininkontinens kan antas svara sämre på larmbehandling än barn med isolerad nattväta.

Bortfall är ett problem vid en majoritet av studierna, vilket sannolikt avspeglar det stora engagemang som larmbehandling kräver av både barnen och deras föräldrar. Detta illustrerar behovet av uppföljning och stöd under och efter behandlingen från utskrivarens sida. Behandlingseffekten kan antas vara högst i välmotiverade, välinformerade familjer.

En kostnadsbedömning (utifrån brittiska förhållanden) ger vid handen att en larmbehandling i 16 veckor torde kosta i genomsnitt en tredjedel av vad desmopressinmedicinering i låg dos under samma period skulle kosta. Till detta kommer att en framgångsrik desmopressinbehandling sällan kan avbrytas efter denna tid utan att följas av återfall.

Behovet av fler studier som jämför larmbehandling och desmopressinbehandling är stort, likaså av studier som söker identifiera prognostiska markörer hos de barn som är lämpade för den ena eller andra behandlingsformen.

II Kommentar

Glazeners och Evans' analys är ett välkommet försök att minska den förvirring som råder kring behandlingen av nattväta. Arbetet försväras av att vi här har att göra med en snärskog av i huvudsak äldre studier av skiftande kvalitet och i vilka enureslarmet jämförs med ett flertal mer eller mindre relevanta terapier. Frågeställningen är kliniskt betydelsefull och besvaras så gott det låter sig göras med de metoder som använts i analysen. Urvalet av studier och värderingen av dessa är adekvata, men fokuseringen på uteslutandet av organisk genes till nattvätan är kanske något övernitisk hos författarna, då det vid isolerad nattväta är mycket sällsynt att det föreligger någon bakomliggande sjukdom.

Sannolikheten att man missat någon väsentlig randomiserad kontrollerad studie av enureslarmbehandling är dock ringa. Det bör emellertid påpekas att till de studier som i sam-

manställningen klassificeras som »awaiting assessment« och alltså ännu inte bedömts hör dels Bradburys studie från 1995, som ger vid handen att larmbehandling kombinerad med desmopressin ger bättre effekt än enbart larmbehandling [17], dels Werrys studie från sextioalet som visar att enureslarmet fungerar mycket bättre än psykoterapi [18].

En pikant detalj i sammanhanget är att det i litteraturgenomgångens referenslista står att läsa följande: »Scholander T [Treating enuresis nocturna by a combination of medicines and conditioning] [German] Lakartidningen 1968;65(46):4552-6«. Kanske var granskningen av artiklar på främmande språk inte den allra bästa?

Trovärdigheten i analysen förstärks under alla omständigheter av att resultaten pekar åt samma håll i de studier som sökt besvara samma fråga. Glazeners och Evans' försiktiga slutsats – att enureslarmet är en effektiv behandling av nattväta – är därför väl underbyggd. Något som ytterligare förstärker denna slutsats, och som författarna själva påpekat, är att det i många av de inkluderade studierna sannolikt ingått även barn med kombinerad dag- och nattväta. Larmet kan på goda grunder antas fungera bäst vid isolerad nattväta.

Vi har alla skäl att anta att litteraturgenomgångens slutsats gäller också svenska barn och är applicerbar i vår kliniska vardag. Enureslarmet försvarar alltså väl, eller bör tillåtas försvara, sin plats som en förstahandsbehandling vid nattväta.

Den andra förstahandsbehandlingen är det antidiuretiska läkemedlet desmopressin (Minirin). Det är olyckligt att så få studier som jämför dessa två behandlingar har genomförts. Här står bara en randomiserad [11] och en orandomiserad [19] studie att finna. Resultaten stämmer dock väl överens med varandra och med den kliniska erfarenheten: behandlingarna är ungefär likvärdiga, men desmopressinet har en snabbare effekt medan larmet har en avsevärt större kurativ potential.

Jämförelser mellan enureslarmet och tricykliska antidepressiva är mindre intressanta, då dessa läkemedel på grund av sin potentiella toxicitet numera bör användas endast i undantagsfall.

Framgång vid behandling av nattväta mäts förstås i torra nätter. Ett problem med den kliniska applicerbarheten av olika studier är att en reduktion av antalet våta nätter ofta räknas som framgång även om den inte är fullständig. Även om varje reduktion av antalet våta nätter självklart är efterlängtat för det sängvätande barnets familj så blir den drabbade själv knappast riktigt nöjd med sin behandling förrän den gör honom eller henne tillförlitligt torr varje natt. Först när man helt kan lita på att man kommer att vakna upp i en torr säng vågar man följa med på skolresor eller sova över hos kamrater.

Det här är något som talar till enureslarmets fördel framför desmopressinbehandling. I studier av enureslarmet används oftast uppnåendet av 14 torra nätter i rad som kriterium på framgång [12, 13, 15], medan de studier av desmopressinbehandling som gjorts ofta definierar terapeutisk respons som en 90-procentig, eller i vissa fall till och med lägre, reduktion av antalet våta nätter [20].

Förenklat uttryckt består den terapeutiska arsenalen idag, sådan den ter sig för läkaren i öppenvården, av två behandlingsalternativ: det antidiuretiska läkemedlet desmopressin (Minirin) eller enureslarmet. Det har märkts en tydlig trend under senare år att läkarkåren i första hand väljer det förra av dessa alternativ. Detta beror sannolikt på att desmopressinbehandling kräver mycket mindre tid, arbete och engagemang än larmet, både för utskrivaren och för de involverade familjerna. Måhända är detta en olycklig utveckling, då larmbehandling har en större kurativ potential och dessutom pris-skillnaden är högst påtaglig, till läkemedlets nackdel.

Huvudproblemet med larmbehandlingen är att den kräver mycket arbete och genererar mycket frustration under be-

II Fakta 1

Grundläggande principer vid behandling med enureslarm

1. Familjen, och inte minst barnet själv, ska vara välmotiverad till behandlingen.
2. Barnet ska veta att de nattliga väckningarna inte är en bestraffning utan ett medel för barnet att lära sig känna när blåsan vill tömma sig.
3. Behandlingen ska vara konsekvent. Inga uppehåll över helgerna!
4. Mamma eller pappa ska sova i samma rum som barnet och hjälpa till med väckningen när larmet ljuder. Det är mer regel än undantag att barnen i början av behandlingen inte själva förmår vakna av ljudsignalen.
5. Det är lämpligt att stänga av larmet efter nattens första kysning, så att barnet (och resten av familjen) inte blir väckt mer än en gång per natt.
6. Har enureslarmet inte givit någon som helst effekt på två månader ska man avbryta behandlingen.
7. Om larmet har en positiv effekt ska man fortsätta behandlingen tills 14 torra nätter i rad har uppnåtts.
8. Vid återfall efter framgångsrik larmbehandling är det lämpligt att försöka igen.

handlingens gång. Detta framgår tydligt av de få studier som har fokuserat på biverkningar och inte bara på behandlingseffekten. Några som helst hälsorisker föreligger dock inte.

II Min konklusion

När läkaren ställs inför ett barn med isolerad nattväta och som har uppnått en ålder då nattvätan börjar upplevas som ett problem, dyker alltid frågan upp: larmet eller Minirin? Mitt råd är då att man inte närmast slentrianmässigt ska skriva ut Minirin utan i stället presentera de två behandlingsalternativen, med deras för- och nackdelar, för familjen och låta denna fatta avgörandet. Detta kommer att leda till att de barn som har störst chans att svara på larmet får just denna behandling och att de övriga slipper drabbas av den stora frustration som en misslyckad larmbehandling innebär.

Enureslarmets något skamfilade rykte beror till stor del på att familjerna ofta fått dålig information om vad de givit sig in på. Lämnar man ut ett enureslarm är det av yttersta vikt att familjen utbildas om hur behandlingen går till och att denna följs upp, åtminstone med ett telefonsamtal efter 1–2 månader. Ska man använda larmet så ska man göra det ordentligt!

Referenser

1. Hägglöf B, Andrén O, Bergström E, Marklund L, Wendelius M. Self-esteem before and after treatment in children with nocturnal enuresis and urinary incontinence. *Scand J Urol Nephrol* 1997;31-(Suppl 183):79-82.
2. Wolfish NM, Pivik RT, Busby KA. Elevated sleep arousal thresholds in enuretic boys: clinical implications. *Acta Paediatr* 1997;86:381-4.
3. Rittig S, Knudsen UB, Nørgaard JP, Pedersen EB, Djurhuus JC. Abnormal diurnal rhythm of plasma vasopressin and urinary output in patients with enuresis. *Am J Physiol* 1989;256:F664-71.
4. Yeung CK, Chiu HN, Sit FK. Bladder dysfunction in children with refractory monosymptomatic primary nocturnal enuresis. *J Urol* 1999;162(3 Pt 2):1049-54, 1054-5.
5. Nevéus T, Läckgren G, Tuveno T, Hetta J, Hjälmås K, Stenberg A. Enuresis – background and treatment. *Scand J Urol Nephrol*

- 2000;202(Suppl 206):1-44.
6. Smith S. Psychological origin and treatment of enuresis. Seattle: University of Washington Press; 1948.
 7. Glazener CM, Evans JH. Alarm interventions for nocturnal enuresis in children (Cochrane Review). In: The Cochrane Library. Issue 1, 2001. Oxford: Update Software.
 8. Butler RJ, Forsythe WI, Robertson J. The body-worn alarm in the treatment of childhood enuresis. *Br J Clin Pract* 1990;44(6):237-41.
 9. Fielding D. The response of day and night wetting children and children who wet only at night to retention control training and the enuresis alarm. *Behav Res Ther* 1980;18:305-17.
 10. Sukhai RN, Mol J, Harris AS. Combined therapy of enuresis alarm and desmopressin in the treatment of nocturnal enuresis. *Eur J Pediatr* 1989;148:465.
 11. Wille S. Comparison of desmopressin and enuresis alarm for nocturnal enuresis. *Arch Dis Child* 1986;61:30-3.
 12. Bollard J, Nettelbeck T. A comparison of dry-bed training and standard urine-alarm conditioning treatment of childhood bedwetting. *Behav Res Ther* 1981;19(3):215-26.
 13. Jehu D, Morgan RT, Turner RK, Jones A. A controlled trial of the treatment of nocturnal enuresis in residential homes for children. *Behav Res Ther* 1977;15(1):1-16.
 14. Sloop EW, Kennedy WA. Institutionalized retarded nocturnal enuretics treated by a conditioning technique. *Am J Mental Deficiency* 1973;77(6):717-21.
 15. Wagner WG, Matthews R. The treatment of nocturnal enuresis: a controlled comparison of two models of urine alarm. *J Dev Behav Pediatr* 1985;6(1):22-6.
 16. Danquah SA. Comparative treatment of nocturnal treatment among Ghanaian school children. *Psychopathologie Africaine* 1975;11(3):363-73.
 17. Bradbury MG, Meadow SR. Combined treatment with enuresis alarm and desmopressin for nocturnal enuresis. *Acta Paediatr* 1995;84:1014-8.
 18. Werry JS, Cohn J. Enuresis – an etiologic and therapeutic study. *J Pediatr* 1965;67:423-31.
 19. Monda JM, Husmann DA. Primary nocturnal enuresis: a comparison among observation, imipramine, desmopressin acetate and bed-wetting alarm systems. *J Urol* 1995;154(2 Pt 2):745-8.
 20. Hjälmås K, Hanson E, Hellström AL, Kruse S, Sillén U. Long-term treatment with desmopressin in children with primary monosymptomatic nocturnal enuresis: an open multicentre study. Swedish Enuresis Trial (SWEET) Group. *Br J Urol* 1998;82(5):704-9.

SUMMARY

Urine alarm defends its position
Comment on a meta-analysis of alarm treatment of
nocturnal enuresis

Tryggve Nevéus

Läkartidningen 2001;98:3212-15

Glazener and Evans have, on behalf of the Cochrane Library, performed a systematic literature review of alarm interventions in nocturnal enuresis. An extensive, and probably exhaustive, search yielded 22 studies which were of sufficient methodological quality and in which the enuresis alarm was included as one treatment alternative. From these studies it could be concluded that 1) the enuresis alarm is clearly more effective than non-treatment, 2) different types of enuresis alarms do not differ significantly as regards efficacy, and 3) alarm treatment is not clearly superior or inferior to pharmacological treatment.

Correspondence: Tryggve Nevéus, Akademiska barnsjukhuset, SE-751 85 Uppsala, Sweden
(tryggve.neveus@kbh.uu.se)



Ska värna de yngre läkarnas intressen

David Eberhard är sedan ett år tillbaka adjungerad redaktör i *Läkartidningens* medicinska redaktion. Davids uppdrag är att företräda de yngre läkarnas perspektiv på det redaktionella materialet, vilket innebär granskning och planering av idédebatt och utbildningsserier som är av speciell relevans för läkare under allmäntjänstgöring och specialistutbildning. Under år 2002 kommer David bl a att ansvara för en artikelserie fokuserad på fortbildningskunskap och utbildning. David svarar ej för fackgranskningen av vetenskapliga artiklar men är rådgivare vid utformningen av *Läkartidningens* vetenskapliga material.

Så här ser David på sitt uppdrag:

»Det är väsentligt att vi betonar vikten av *Läkartidningen* som en utbildningsresurs och därför måste vi fokusera på material som har ett stort utbildningsvärde. För att göra det, är det nödvändigt att bredda vissa artiklar och göra dem mer lättlästa även för de doktorer som inte är specialintresserade.

Det kan vara svårt, för samtidigt är det ju oerhört viktigt att man behåller artiklar som faktiskt tillför de specialkunniga nya aspekter«.

David är redaktör för Sylfs tidning *Moderna Läkare* och har genomgått journalistutbildning innan han påbörjade sina läkarstudier. Han gör för tillfället sin randutbildning på neurologiska kliniken vid Karolinska sjukhuset i Stockholm. Sin ordinarie tjänst har han som ST-läkare i psykiatri vid S:t Görans sjukhus i Stockholm och han blir färdig våren 2002. David är 34 år och har två barn.

För *Läkartidningens*
medicinska redaktion
Josef Milerad
medicinsk chefredaktör

Välkommen SBU-rapport visar att det finns effektiva metoder mot missbruk

På vissa områden har tiden dock sprungit ifrån rapporten

II Det finns idag en utbredd behandlingspessimism på missbruksområdet, som i kombination med ansträngd offentlig ekonomi lett till nedskärningar inom missbrukarvården. En ny rapport från SBU visar tydligt att uppgivenheten inte är motiverad. Rapporten är därför mycket välkommen. De goda nyheterna först: Det finns metoder med dokumenterad effektivitet på missbruksområdet. Effektiviteten är inte sådan att man kan slå sig till ro, men inte heller sämre än vid många andra folksjukdomar, till exempel hypertoni eller diabetes. De dåliga nyheterna, från den kliniska verkligheten: Det är inte de effektiva metoderna som erbjuds flertalet missbrukare.

Innan beroende utvecklats – mini-intervention

En alltför oanvänd kunskap lyfts fram av SBU-rapporten. Enkla screening-insatser i kombination med enstaka rådgivningssessioner har utmärkta effekter för att reducera ohälsosam alkoholkonsumtion. Därmed sjunker sjuklighet och sjukvårdskonsumtion. Det är en utmaning för ett pressat sjukvårdssystem att implementera denna metodik, vars vinster ligger en bit bort i tiden, samtidigt som en störtflod av akuta problem måste hanteras.

Parallellt med entusiasmen för mini-intervention finns anledning att höja ett varnande finger. Metodiken är billig. För rätt patienter – högkonsumenter med god social funktion, och utan etablerat beroende – är den också mycket kostnadseffektiv. Sparivriga sjukvårdspolitiker kan dock lätt falla för frestelsen att lansera metodiken som alternativ till dyrare insatser för personer med höggradigt beroende. I värsta fall kan budskapet fungera som täckmantel för etiskt oförsvärbara nerdragningar, något som redan inträffat i Stockholmsregionen.

Långsiktig behandling – för patient eller behandlare?

Den vanligaste behandlingen för alkoholberoende i Sverige är ostrukturerad stödkontakt, ofta kombinerad med oövervakad disulfirammedicinering. Det är sedan tidigare välkänt att dessa behandlingsmodaliteter saknar effekt [1], vilket bekräftar av SBU-rapporten. Psykodynamisk psykoterapi, som har hög status bland behandlare på fältet, saknar även den dokumenterad effekt, och är dessutom rekorddyr. Samtidigt finns

Evidensbaserad medicin

Se även artikeln på sidan 3906 i detta nummer.

dokumentation för att relativt enkla beteendeorienterade interventioner är verksamma, än mera så i kombination med modern läkemedelsbehandling, till exempel opiatreceptor-antagonisten naltrexon, eller NMDA-/spermidinantagonisten akamprosatsat. En försvinnande minoritet av svenska alkoholberoende patienter erhåller sådan behandling.

Narkomanibehandling – evidens eller politisk korrekthet?

Ett vetenskapligt gediget och kliniskt relevant kapitel i SBU-rapporten redovisar kunskapsläget vad gäller psykosociala behandlingsmetoder vid narkomani. De övergripande slutsatserna är tydliga, överensstämmer med klinisk erfarenhet och

Det är inte de effektiva metoderna som erbjuds flertalet missbrukare.

går i samma riktning som alkoholdelen. Det är strukturerade, beteendeorienterade interventioner med tydligt fokus på missbruket som är effektiva för retention i behandling och minskning av missbruk.

Ett centralt faktum återfinns i sammanfattningen endast som en lätt förbisedd mening: »Samtliga studier har dock gjorts med patienter som också behandlas med metadon.« Förmodligen anses detta ha för kontroversiella implikationer för att ges det utrymme som motsvarar iakttagelsens dignitet. Innebörden är nämligen att inte ens de bästa psykosociala behandlingsmetoderna har dokumenterad effekt för heroin-

beroende patienter om de inte kombineras med verksam läkemedelsbehandling, det vill säga metadon eller liknande medel. Detta är välkänt, och beror på att patienterna i annat fall helt enkelt inte kvarstannar i behandling [2]. Färre än 10 procent av svenska heroinister erhåller läkemedelsassisterad behandling. Den politiskt omhuldade dogmen är att läkemedelsbehandling skall undvikas, och när den ändå används endast utgöra ett komplement till psykosocial behandling. Evidensbasen visar att det i själva verket är tvärtom. Det skall bli intressant att se hur starkt försvurna åt evidensbaserade principer vi är när det verkligen gäller. Med tanke på den höga mortaliteten vid heroinberoende är frågan inte främst akademisk.

Läkemedelsbehandling vid narkomani behandlas i ett mycket svagare kapitel. Det överväldigande stödet för metadonbehandlingens effektivitet vid heroinnarkomani framträder förstås tydligt. Initiala studier med den partiella agonisten buprenorfin refereras, och indikerar att även detta preparat är verksamt, en slutsats som stöds av senare tillkomna studier [3]. En diskussion av plasmanivåernas betydelse för den relativa effektiviteten av olika medel saknas. Kapitlet andas optimism vad gäller användning av naltrexon vid heroinnarkomani. Efter 30 års försök har dessa förhoppningar övergivits av de flesta i fältet, annat än för speciella patientgrupper med främst unga patienter tidigt i beroendutvecklingen.

Slutsatserna för behandling av opiatabstinens baseras på ett fåtal gamla studier och tveksamt relevanta utfallsmått. Klonidin är lika effektivt som korstoleranta medel (opiater) för att hämma vegetativa abstinenssymtom, men vad som plågar patienten och gör att avgiftningen avbryts är inte detta, utan de psykiska symtomen, mot vilka klonidin saknar effekt. Tiden har här sprungit ifrån SBU-rapporten. Sedan buprenorfin registrerats i Sverige använder ingen av landets ledande narkomanvårdsenheter klonidin som förstahandsmedel vid behandling av opiatabstinens.

Svenskt forskningscentrum bör upprättas

Förutom banbrytande men två decennier gamla metadonarbeten av professor Gunne och medarbetare [4] är nästan inga av studierna SBU-rapporten förlitar sig på svenska. På centralstimulantiaområdet får vi därför god kunskap om vad som fungerar för kokainister, men inte för amfetaminpatienter, som helt dominerar i vårt land. Varför saknas god svensk forskning, och vad kan vi göra åt det? För att bli framgångsrik kan man lära av dem som redan har framgång. Vi borde snarast upprätta ett nationellt forskningsinstitut på missbruksområdet, med amerikanska NIAAA och NIDA som förebild. Förutom att initiera, finansiera och koordinera högkvalitativa studier skulle ett sådant kunna fylla en långsiktigt lika viktig uppgift, att bli ett kunskapscentrum som utbildar behandlare, allmänhet och politiker.

Då kanske avståndet mellan SBU-rapportens kunskapsläge och den kliniska verkligheten på sikt kunde bli mindre skriande än idag.

Referenser

1. Hester RK, Miller WR, editors. Alcoholism treatment approaches: Effective alternatives. Needham Heights: Allyn and Bacon, 1995.
2. Sees KL, Delucchi KL, Masson C, Rosen A, Clark HW, Robillard H, et al. Methadone maintenance vs 180-day psychosocially enriched detoxification for treatment of opioid dependence: a randomized controlled trial. JAMA 2000;283(10):1303-10.
3. Johnson RE, Chutuape MA, Strain EC, Walsh SL, Stitzer ML, Bigelow GE. A comparison of levomethadyl acetate, buprenorphine, and methadone for opioid dependence. N Engl J Med 2000;343(18):1290-7.
4. Gunne LM, Grönbladh L. The Swedish methadone maintenance program: a controlled study. Drug and Alcohol Dependence 1981;7(3):249-56.

Särtryck Läkartidningen

När konsensus saknas om hur läkaren bör behandla, spelar den beprövade erfarenheten stor roll. Det 48-sidiga häftet innehåller 32 korta, praktiskt inriktade artiklar med anknytning till vårdens vardag och vänder sig till alla kliniskt verksamma läkare. Förutom diagnostik med terapi speglas goda exempel på prevention, ledningsfrågor och administration.

Priset är 55 kr.

Enligt min erfarenhet



Beställer härmed ex
av "Enligt min erfarenhet"

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till Läkartidningen
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Cochrane-rapport om akupunktur vid huvudvärk

Referat och kommentarer av en systematisk översikt

|| Akupunktur är en mer än 2 000 år gammal terapeutisk metod där man genom att sticka in tunna, solida nålar på olika ställen i kroppen uppges kunna lindra smärta och en del andra symtom. Metoden är en del av den så kallade traditionella kinesiska medicinen (TCM). Enligt TCM påverkar och balanserar akupunktur energiflöden i kroppen medan vi med modern fysiologi betraktar akupunktur som en form av somatisk afferent sensorisk stimulering. Denna stimulering leder till lokala läkeeffekter ute i vävnaden samt påverkar olika endogena smärthämningssystem. En översiktsartikel har nyligen publicerats i Läkartidningen [1], där mer kan läsas om dess bakgrund, användning och troliga mekanismer.

|| Evidensbaserad akupunktur

Även akupunktur har det senaste decenniet studerats enligt de metoder som används inom evidensbaserad medicin, så kallad evidensbaserad akupunktur. Det är emellertid stora problem med att summera akupunkturstudier, eftersom såväl sättet att utföra akupunktur som studiernas uppläggning är ytterst skiftande. I kliniken letar man efter så bra och variabel »dos« av akupunktur som möjligt. Denna dos beror på faktorer som: antalet använda akupunkturpunkter, vilka punkter som har använts, hur tjocka nålarna var, hur mycket man stimulerade nålarna manuellt (eller elektriskt), hur lång tid nålarna satt kvar, vilket intervall man hade mellan behandlingarna, hur många gånger akupunktur gavs totalt etc. Denna akupunkturs läkekonst innebär att man i hög grad individualiserar sina behandlingar – något som gör klinisk forskning besvärlig. Man kan säga att överlag föreligger problem att få akupunkturbaserad evidens när man gör dessa översikter.

Huvudvärk, såväl migrän som spänningstypen, är en mycket vanlig indikation för akupunktur, även om konklusiva studier hittills saknas vad gäller effektiviteten. För att råda bot på denna brist har nu en kvalitativ systematisk översikt publicerats från Cochrane Collaboration [2].

Metodologiska problem

Tre metodologiska huvudproblem finns när det gäller akupunkturstudier: dubbelblindningen, placeboalet och behandlingsupplägningen [3, 4]. Sammanlagt leder de två förstnämnda problemen till att studier ofta erhåller tämligen

SAMMANFATTAT

En kvalitativ systematisk översikt har publicerats från Cochrane Collaboration, där man har analyserat effekten av akupunktur vid olika former av huvudvärk.

Studiernas skiftande kvalitet och uppläggning gjorde att en kvantitativ metaanalys inte kunde göras.

Majoriteten av studier som jämförde riktig akupunktur med shamakupunktur vid migrän och spänningshuvudvärk visade åtminstone en trend till akupunkturs fördel.

Det finns inga studier som endast gör jämförelser med obehandlade kontroller.

Det finns inga studier som har undersökt vilken »dos« av akupunktur som borde vara bäst.

Det är stor brist på långtidsuppföljningar.

Inom området föreligger fortfarande alltför stora oklarheter för att med EBM-metodik fastställa effekt.

Evidensbaserad medicin

låga metodpoäng på sin vetenskapliga styrka. Det tredje problemet leder till att en värdering av akupunkturs utförande ger låga poäng. Detta tredje problem är i många systematiska översikter inte värderat alls. Emellertid är det i nu aktuell översikt värderat hur bra akupunkturs utförande blivit jämfört med i kliniken vanligen använd teknik.

Egentliga dubbelblinda studier kan för det mesta inte utföras när man undersöker metoder som akupunktur, kirurgi, fysikalisk terapi eller psykologi. Detta sammanhänger bland

annat med att behandlaren helst bör vara medveten om vad han/hon gör. Dessutom kräver en del metoder att patienten samarbetar för att utförandet ska bli optimalt. Som exempel kan nämnas att det anses mycket viktigt att patienten erfar en speciell »nålkänsla« (de-qi), då nålen kommit i »rätt« läge, för att akupunktur ska ha fullgod effekt.

Det bästa som normalt kan åstadkommas när det gäller just blindningen är att patienten »blindas«, och att utvärderaren är »blind«. Vissa väljer att kalla dylika studier dubbelblinda, men som regel kallas de enkelblindade med oberoende utvärderare.

Placebokontrollen i akupunkturstudier har debatterats mycket [1, 3, 4]. Man kan säga att många av de tidigare studierna egentligen bara har visat att TCM-teorin inte är korrekt, eftersom man inte har funnit signifikanta skillnader i effekt mellan att sätta akupunktur nålar i traditionella akupunkturpunkter eller i andra närbelägna icke-akupunkturpunkter. Eftersom vi anser akupunktur vara en form av sensorisk afferent stimulering är det troligen bara en fråga om att man stimulerat punkter med olika rik nervös innervation. Slutsatserna har blivit att akupunktur inte har någon effekt mer än placebo. Konklusionen i många studier hade precis lika gärna kunnat vara att det inte förelåg någon statistisk skillnad mellan att med nålar stimulera traditionella eller icke-traditionella punkter, speciellt eftersom det i de allra flesta fall visade sig att nålstimuleringen hade positiva effekter jämfört med baslinjeregistreringar. En annan form av använd placebo har varit så kallad minimal akupunktur som innebär en minimal ytstimulering, där nålarna sätts utanför traditionella akupunkturpunkter, endast införda 1–2 mm i skinnet, och stimulerade extremt lätt. Dock har det visats att även denna minimala stimulering har en del kliniskt positiva effekter [5–8]. Alltså, yttlig eller minimal akupunktur bör betraktas som en annan och svagare form av sensorisk stimulering men inte som en totalt inert placebobehandling.

Under de senaste åren har man utvecklat en ny typ av placebo nålar, där själva nålen (som är helt planslipad) glider upp i handtaget på nålen när man försöker få den igenom huden [9, 10]. Man sätter nålen (riktig eller placebo) genom ett membran och en liten plasthållare som gör att man inte kan se om nålen penetrerar skinnet eller inte. Dock är att märka att även denna teknik leder till en sensorisk afferens eftersom en del försökspersoner faktiskt erfar nålkänsla även av denna lilla stimulering. Dessutom kan till metodens nackdel räknas att utövaren inte kan söka de-qi på vanligt vis, eftersom man då bryter »blindningen«. Alltså förändrar själva experiment-situationen betingelserna för metoden även i detta fall.

II Litteraturgranskning

Denna artikel summerar och kommenterar en kvalitativ systematisk litteraturoversikt från Cochrane Collaboration, vars syfte var att besvara frågan huruvida akupunktur har effekt mot »idiopatisk huvudvärk«. Mer detaljerat var målet att utvärdera om akupunktur

- är bättre än ingen behandling,
- är bättre än shamakupunktur (placebo),
- är likvärdig med andra vanliga behandlingar mot huvudvärk.

Material och metoder

Sökstrategin omfattade Medline (1966 – april 2000), Embase (1989 – april 2000), Cochrane controlled trials register (2000, ver 1) samt databasen Cochrane field for complementary medicine. Man kontaktade också forskare inom området för att få reda på eventuellt opublicerade studier inklusive pågående studier. Förutom detta gjordes manuell genomgång av referenslistorna i alla granskade artiklar för att om möjligt fin-

na fler studier. Sökorden var »acupuncture« eller »acupressure« kombinerade med »headache« eller »migraine«.

Urvalskriteriet för de kliniska studierna var att de skulle vara randomiserade eller kvasirandomiserade och jämföra akupunktur med någon typ av kontrollintervention vid behandling av idiopatisk huvudvärk. Till idiopatisk huvudvärk räknade man: migrän, spänningshuvudvärk, Hortons huvudvärk samt oprecist klassificerat som kronisk eller recidiverande primär huvudvärk. Ansiktssmärta ingick inte i översikten.

Interventionen skulle bestå av nålakupunktur med användande av traditionella akupunkturpunkter, »smärtpunkter« eller triggerpunkter. Dessutom ingick några studier (N=3) där man använt annan stimulering av akupunkturpunkter, »laserakupunktur« och TENS.

Kontrollgruppen skulle antingen inte ha erhållit någon behandling eller fått placebo (shamakupunktur)-behandling eller annan behandling.

Man identifierade totalt 104 studier med originaldata som handlade om akupunktur och huvudvärk. Två olika granskare exkluderade direkt 58 av studierna eftersom de inte var kontrollerade. De övriga 46 detaljstuderades av minst två granskare, och efter deras samlade bedömningar uteslöts ytterligare 18 studier som inte uppfyllde inklusionskriterierna eller var oklart rapporterade. Kvar blir alltså 28 studier, men hela rapporten bygger på analys av 26 studier, vilket troligen beror på att ett par studier var fortgående. Det framgår dock inte helt klart i rapporten varför dessa två uteslutits.

Minst två oberoende granskare extraherade data ur studierna med hjälp av ett testat formulär. Vid behov kontaktades författare för att erhålla mer information.

Den metodologiska kvaliteten på studierna värderades också av två oberoende granskare. Man använde två olika skalor för ändamålet. Den ena, Jadadskalan [11], ger poäng mellan 0 och 5, där poäng från 3 och uppåt anses som högkvalitativa studier. Man bedömer där randomiseringen, dubbelblindningen samt bortfallen. Den andra skalan (Internal validity scale, IVS) gör en liknande bedömning, men på ett sätt som anges vara mer tillämplig för studier inom kompletterande medicinska området [12].

Kvaliteten på utförd akupunktur värderades av en granskare, specialiserad inom såväl akupunktur som huvudvärk. Denne utförde värderingarna helt blint genom att han endast fick läsa metodavsnitten i studierna, där alla andra avsnitt såväl som författarnamn och källa hade tagits bort. Han värderade på en skala från 0 till 100 procent hur korrekt han ansåg att akupunkturen var. Siffran 100 procent betydde att granskaren var till 100 procent övertygad om att akupunkturen var korrekt utförd, och att denne skulle ha gjort precis likadant själv. I sju av studierna kunde inte denna värdering göras eftersom akupunkturtekniken var alltför bristfälligt rapporterad.

Totalt granskades alltså 26 studier, och dessa innehöll sammanlagt 1 151 patienter.

16 av studierna gällde migrän, 6 gällde spänningstypen och 4 innehöll blandade diagnoser.

17 av studierna var placebokontrollerade, där 11 använde shamakupunktur (andra punkter än akupunkturpunkter alternativt ytliga nålar) och 6 använde andra procedurer som inte innebar penetration av huden.

4 av studierna jämförde akupunktur med farmaka.

De övriga 5 jämförde akupunktur med sjukgymnastik, relaxation, traditionella örter samt beteendeprogram.

3 av studierna gällde inte vanlig nålakupunktur utan laser (N=2) eller TENS med plattorna applicerade över akupunkturpunkter (N=1).

I genomsnitt hade man behandlat patienterna åtta gånger under åtta veckor. 21 av studierna hade följt upp patienterna, ■

vilket skett i genomsnitt 26 veckor efter avslutad behandlingsserie.

I 8 av studierna valdes akupunkturpunkter helt i enlighet med TCM medan man i 7 av studierna angav andra skäl. I 10 studier gav man ingen förklaring till sitt punktval. En studie använde endast lokalt ömma punkter.

Basala data om kvalifikationen och erfarenheten hos den/de akupunkturutövande gavs bara i 10 av studierna.

Nålkänslan, de-qi, den utstrålade känslan som anses mycket viktig för fullgod effekt av akupunkturstimulering, angavs sällan i studierna.

Utvärderingsmått var oftast huvudvärksfrekvens och smärtintensitet, men många studier rapporterade detta mycket dåligt. Huvudvärksdagbok användes i 18 av studierna. Data i huvudvärksdagböckerna använde man för att få en global skattning om studien saknade uppgift om detta. I dessa fall definierades minst 33 procents förbättring på frekvens eller huvudvärksindex, jämfört med baslinjesregistrering, som förbättrat.

Otillräcklig noggrannhet vid datapresentationen samt olikheterna i studierna gjorde en formell kvantitativ metaanalys meningslös. I stället värderade minst två oberoende granskare studierna till någon av följande kategorier vad gällde global effekt: kontrollgruppen bättre än akupunkturgruppen (signifikant), kontrollgruppen bättre än akupunkturgruppen (trend), ingen skillnad mellan grupperna; akupunkturgruppen bättre än kontrollgruppen (trend); akupunkturgruppen bättre än kontrollgruppen (signifikant). Vid oenighet (N=2) gällde den minst fördelaktiga skattningen för akupunkturgruppen.

Man anger att majoriteten av studier hade rent metodologiska problem eller problem med rapporteringen av resultaten. Text var medianvärdet för skattningarna enligt Jadadskalan 1,5. I flertalet studier var bortfallsrapporteringen bristfällig.

Resultat

Det var totalt 16 studier innehållande patienter med migrän eller spänningstypen av huvudvärk, där man jämförde riktig akupunktur med shamakupunktur. I 8 av dessa 16 studier var riktig akupunktur signifikant överlägsen placebo, hos 4 av de 16 fanns en trend i samma riktning, och i 2 studier fanns ingen skillnad mellan grupperna. 2 av dessa 16 studier gick inte att tolka. Framför allt var det bland migränstudierna som man såg uppmanande resultat.

De 10 studierna som jämförde akupunktur med annan behandling gav motstridiga och svårtolkade resultat.

När det gällde långtidseffekter fanns 13 studier som hade uppföljning minst sex månader efter terapin. De flesta hade hög frekvens bortfall, dålig resultatredovisning eller allmänt dålig validitet. 2 av de mer rigorösa studierna visade att effekterna kunde stå sig mer än sex månader efter avslutad behandling.

Den metodologiska kvaliteten på studierna visade att det fanns 5 tolkbara studier (migrän N=3, spänningstypen N=2) med 3 poäng eller mer enligt Jadadskalan som jämförde nålakupunktur mot placebo. Av dessa 5 var 1 signifikant positiv för akupunkturgruppen, 2 trendpositiva och 1 utan skillnad mellan grupperna, och 1 var trendnegativ.

När man i stället utgick från kvaliteten på akupunkturbehandlingarna fanns det 9 tolkbara med poäng på 70 procent eller mer. Av dessa var 6 signifikant positiva och 3 trendpositiva, medan ingen av dessa visade »ingen skillnad« eller bättre resultat hos kontrollgrupperna. Även här är det flest migränstudier som är positiva.

Diskussion och konklusion

Totalt ger översikten ett visst stöd för att akupunktur är av värde vid behandling av idiopatisk huvudvärk. Emellertid är inte bevisens kvalitet och mängd helt övertygande.

Studierna var i hög grad heterogena vad gällde patienter,

akupunkturs utförande, kontrollinterventionen samt även när det gällde effektmåten. De flesta studierna hade metodologiska och/eller rapporteringsproblem.

Majoriteten av studier som jämförde sham och riktig akupunktur vid migrän och spänningstypen visade åtminstone en trend i riktning mot akupunkturs fördel. Dock betonas det att de flesta studierna var små och hade olika metodologiska tillkortakommanden.

Man konstaterade vidare att det faktiskt inte finns bevis för att akupunktur är bättre än ingen behandling likaväl som det inte finns jämförelser mellan akupunktur och vanliga andra behandlingar. Man konstaterade, som vanligt, att det behövs välplanerade, storskaliga studier för att få reda på effekt och kostnad-nytterelationen vid akupunktur under verkliga omständigheter.

II Diskussion kring Cochranerapporten

Översikten är, som alltid när det gäller Cochranerapporter, ambitiös och verkar överlag noggrant utförd. Emellertid finns det inbyggda problem i kvalitativa översikter, eftersom de i viss mån bygger på subjektiva bedömningar som leder till att »röstning« förekommer. Att inte en riktig metaanalys gjorts är beklagligt men fullt förståeligt, eftersom stora olikheter förelåg mellan de ingående studierna. Att göra en sammanräkning av studier där vissa är vetenskapligt metodologiskt tveklaktiga och en del är akupunkturmässigt »okliniskt« upplagda stöter naturligtvis på problem.

Intressant i den aktuella översikten är att man värderar akupunkturutförandet och inte bara den rent formella vetenskapliga uppläggningsen. Resultaten av värderingarna framgår ovan. Vad som inte nämns i rapporten, men som man lätt själv kan få fram, är följande (gällande nålakupunktur): Om man kombinerar dem med god kvalitet enligt Jadadskalan (3 poäng eller mer) med dem med god akupunkturkvalitet (70 procent eller mer) så finns 2 studier, varav 1 är positiv och 1 trendpositiv. Mer intressant är kanske följande: Av tolkbara med mindre än 70 procent på akupunkturkvaliteten (N=7) var 2 trendpositiva medan 5 visade »ingen signifikant skillnad«. Således tycks det som att värderingen av akupunkturkvaliteten är av stor vikt när man ska bedöma värdet av en studie eller göra en systematisk översikt.

Översikten har emellertid ett par problem som man bör beakta vid tolkningen av den. Man har blandat olika diagnoser vid ett flertal analyser. Man har slagit ihop olika stimuleringsätt som inte alls har med akupunktur att göra, laser (N=2) och TENS (N=1). Att kombinera laserstudier och akupunkturstudier är inte alls bra, eftersom det rör sig om helt olika kategorier av metoder (behandling med en ljusstråle analyseras tillsammans med mekanisk nålstimulering). Även TENS är en annan form av stimulering, eftersom denna innebär aktivering av hudafferenter medan akupunktur som regel innebär aktivering av muskelafferenter. Ytterligare en oklarhet i översikten är att en del av studierna anges som »ej tolkbara« och några som »oklara«. Trots detta har de emellertid tagits med i flera av analyserna.

Författarens konklusion

Det enda som översikten kunde konkludera var att majoriteten av studier som jämförde sham och riktig akupunktur visade åtminstone en trend till akupunkturs fördel. Man konstaterade vidare att det faktiskt inte finns bevis för att akupunktur är bättre än ingen behandling (brist på studier) likaväl som det inte finns jämförelser mellan akupunktur och vanliga andra behandlingar. Dessutom saknas nästan helt långtidsstudier.

När det gäller kliniska akupunkturstudier har man extensivt diskuterat dubbelblindningsteknik och placebo-teknik. Med hänsyn till den erfarenhet som nu finns av kliniska akupunkturstudier sedan 1970-talet, och de tolkningssvårigheter

som så gott som alltid föreligger, borde man gå över till ett mer strukturerat angreppssätt för att få reda på dess kliniska effekt. En rimlig primär ansats vore att starta med att se efter om akupunktur i klinisk användning har någon effekt alls vid olika diagnoser. Detta inte minst eftersom man troligen snabbt skulle kunna utesluta en del diagnoser, då effekten skulle visa sig vara alltför svag jämfört med modern annan behandling för att omsättas i större skala i praktiken. Detta skulle kunna göras genom noggranna analyser där kontrollen är baslinjeregistreringar och/eller obehandlad grupp. Man borde betrakta dylika studier som vad som i farmakaforskning kallas fas 2-studier, dvs studier vars avsikt är att komma fram till vilken »dosering« av behandlingen som är bäst [13].

När man väl vet vilken dosering som är bäst kan man göra formella placebostudier och studier där man jämför med andra behandlingsmetoder. Jämför det absurda i att göra läkemedelsförsök när man inte har definierat effektiv dos utan behandlar med kanske en tiondel av den som är verkningsfull!

Slutsatsen i översikten överensstämmer i stort sett med andra översikter av akupunkturs effekt vid andra smärttillstånd. De flesta författare kommer fram till samma konklusion som meddelas här: Det behövs välplanerade, storskaliga studier med långtidsuppföljning under kliniskt verkliga omständigheter. En slutsats som är rimlig, dock med tillägget ovan, är att det behövs doseringsstudier innan man ger sig i kast med dessa storskaliga »slutgiltiga« studier.

Inom området behövs definitivt inte fler metaanalyser (eller systematiska översikter) utan fler på olika sätt upplagda kliniska långtidsstudier.

Referenser

1. Carlsson C. Akupunktur som behandlingsmetod idag – bakgrund, klinisk användning, mekanismer. *Läkartidningen* 2001; In Press.
2. Melchart D, Linde K, Fischer P, Berman B, White A, Vickers A, et al. Acupuncture for idiopathic headache (Cochrane Review). In: *The Cochrane library*, Issue 1, 2001. Oxford: Update Software, 2001.
3. Vincent CA, Richardson PH. The evaluation of therapeutic acupuncture: concepts and methods. *Pain* 1986;24:1-13.
4. Lewith GT, Vincent CA. Evaluation of the clinical effects of acupuncture. A problem reassessed and a framework for future research. *Pain Forum* 1995;4:29-39.
5. Macdonald AJR, Macrae KD, Master BR, Rubin AP. Superficial acupuncture in the relief of chronic low back pain. A placebo-controlled randomised trial. *Ann R Coll Surg Engl* 1983;65:44-6.
6. Kelleher CJ, Filshie J, Burton G, Khullar V, Cardozo LD. Acupuncture and the treatment of irritative bladder symptoms. *Acupuncture in medicine* 1994;12:9-12.
7. Blom M, Dawidson I, Lundeberg T, Angmar-Månsson B. Effects on local blood flux of acupuncture stimulation used to treat xerostomia in patients suffering from Sjögren's syndrome. *J Oral Rehabil* 1993;20: 541-548.
8. Wyon Y, Lindgren R, Lundeberg T, Hammar M. Effects of acupuncture on climacteric vasomotor symptoms, quality of life, and urinary excretion of neuropeptides among postmenopausal woman. *Menopause* 1995;2:3-12.
9. Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. *Lancet* 1998;352:992.
10. Kleinhenz J, Streitberger K, Windeler J, Gussbacher A, Mavridis G, Martin E. Randomised clinical trial comparing the effects of acupuncture and a newly designed placebo needle in rotator cuff tendinitis. *Pain* 1999;83:235-41.
11. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds M, Gaghan DJ, et al. Assessing the quality of reports of randomised clinical trials: is blinding necessary? *Controlled Clinical Trials* 1996;17:1-12.
12. Linde K, Clausius N, Ramirez G, Melchart D, Eitel F, Hedges LV, et al. Are the clinical effects of homeopathy placebo effects? A meta-analysis of placebo-controlled trials [published erratum appears in *Lancet* 1998;351:220]. *Lancet* 1997;350:834-43.
13. Spilker B. *Guide to clinical trials*. New York: Raven Press, 1991.

Hur man skriver ett vetenskapligt arbete



– och får det publicerat!

Läkartidningen och **BMJ** (British Medical Journal) arrangerar i samarbete med **Karolinska institutet** två endagskurser för medicinska skribenter.

Kurserna vänder sig i första hand till forskarstuderande med minst ett publicerat eget arbete.

Kursen presenterar hur man på ett effektivt sätt omsätter tankar och koncept till en välskrivna uppsats – hur man organiserar och analyserar sitt material, formulerar slutsatser och hur man korresponderar med redaktioner och redaktörer enligt »spelets regler«.

Kursledare är Tim Albert, vetenskapsjournalist på **BMJ** samt Harvey Markovitch, chefredaktör på **Archives of Disease in Childhood**. Kurserna har gått flera år i Storbritannien och rönt stor uppskattning.

Kursdatum måndagen den 3:e december och tisdagen den 4:e december 2001.

Kursavgift 2 000 kr. Karolinska institutet betalar kursavgiften för doktorander registrerade vid KI.

Så här anmäler Du Dig. Ingen särskild anmälningsblankett behövs. Skicka in fullständiga adressuppgifter, registreringsdatum till forskarutbildningen, ett CV på högst 300 ord samt en kort motivering varför Du vill gå kursen.

Forskarstuderande vid KI skall skicka ansökan till Mikael Holst KI (Mikael.Holst@kbh.ki.se)

Forskarstuderande vid andra medicinska högskolor skall skicka ansökan till Josef Milerad Läkartidningen (Josef.Milerad@lakartidningen.se)

Din anmälan måste vara hos oss senast den 7 november.

Läkartidningen vill genom kursen stimulera intresset för att skriva vetenskap på svenska och kommer att betala kursavgiften för 5 deltagare intresserade av att medverka som skribenter av vetenskapliga referat.

Välkommen med Din anmälan
– antalet platser är begränsat!

Josef Milerad
docent, medicinsk chefredaktör

Ny metaanalys av albumin inom intensivvården

Ökning av dödligheten inte statistiskt säkerställd



Albumin – betydligt dyrare och kanske farligare än syntetiska kolloider.

II Intravenös infusion av albumin har på senare tid framstått som en tvivelaktig behandling av hypovolemi, hypoalbuminemi och brännskador inom intensivvården. Orsaken är främst en systematisk litteraturöversikt från the Cochrane Collaboration, som tidigare kommenterats i Läkartidningen [1]. Denna översikt analyserade överlevnaden hos drygt 1 400 patienter från 30 studier som randomiserats till att erhålla eller inte erhålla albumin. Resultatet visade att prognosen försämras när albumin

ges. Den relativa risken för död var 1,46 när albumin givits för att behandla hypovolemi, 1,69 om indikationen var hypoalbuminemi, och 2,40 om albumin givits vid brännskador [2, 3].

En livlig debatt uppstod snart kring denna Cochraneöversikt. Många läkare kände tveksamhet inför värdet av en metaanalys som bygger på variation i en så liten detalj som valet av infusionsvätska i en så pass komplex klinisk situation som intensivvård. Kritikerna framförde dessutom att kvaliteten på de ingående studierna var mycket varierande. Volymterapi gavs ofta utan att doseringen styrts efter hemodynamiska kriterier. Inget alternativ till albumin gick heller att urskilja, då kontrollgruppen kunnat få olika alternativa behandlingar, inklusive ingen behandling alls.

Samtidigt som kritiken mot översikten tätnade påverkade dess slutsatser läkarnas handlande. I England minskade användningen av albumin med 40 procent månaderna efter översiktens publicering [4]. Även andra studier pekade åt samma håll. Två metaanalyser från denna tid visade att volymsubstitution med kolloid vätska är associerad med en högre dödlighet än den efter tillförsel av kristalloid vätska [5, 6]. De tycktes alltså ge stöd för hypotesen att albumin är mindre

SAMMANFATTAT

Intravenös infusion av albumin från blodplasma har i en tidigare litteraturöversikt från Cochranegruppen visats öka mortaliteten inom intensivvården.

I en ny metaanalys ökade dödligheten med 11 procent efter användning av albumin.

Denna ökning försvann dock när man endast analyserade studier som uppfyller minst ett av fyra krav på god kvalitet.

Evidensbaserad medicin

lämpligt att använda än kristalloid vätska som volymsubstitution till svårt sjuka människor.

Nya perspektiv i ny översikt

I augustinumret av *Annals of Internal Medicine* publiceras en ny systematisk litteraturöversikt över riskerna med albuminbehandling [7]. Analysen omfattar 3 504 patienter från 55 studier, vilket är betydligt fler än i Cochraneöversikten. Kriterierna för exklusion har varit få. Man har i stort sett endast krävt att studierna skall innefatta randomisering mellan albumin och kristalloid vätska samt att uppgifter om överlevnad har gått att uppbringa. Nya perspektiv anläggs också. Författarna studerar nämligen utfallet med avseende på kvaliteten i studiernas upplägg.

Resultatet visar en allmän tendens till ökad mortalitet efter albuminbehandling, men skillnaden mot kristalloid behandling är ej statistiskt signifikant. Den relativa risken (RR) för död är 1,11 med 95 procents konfidensintervall på 0,95–1,28. Riskökningen är inte heller signifikant för någon klinisk undergrupp, även om det ibland är snubblande nära: den relativa risken är 1,12 för kirurgi och trauma; 1,76 för brännskador; 1,59 för hypoalbuminemi; 0,93 för ascites; 1,19 för späd-

barn. Men stora skillnader i risk föreligger beroende på hur studierna genomförts.

Fyra kvalitetsaspekter

Författarna har utvärderat effekten av de fyra kvalitetsaspekterna blindning, mortalitet som slutmättningspunkt, »crossover» samt antalet studerade patienter. Risken med albumin visade sig vara lägre ju bättre kvalitet studien hade. Vi skall nu titta närmare på dessa effekter.

Bara fem studier var blindade, men dessa uppvisade snarast en minskad risk (RR 0,73). De 17 studier där död var endpoint indikerade heller inte någon riskökning vid användning av albumin (RR 1,00). Orsaken till detta tolkar författarna som att övriga studier möjligen har saknat adekvata metoder för registrering av dödsfall. Uppföljningstiden kan också ha varit för kort. Minst lika viktig torde förekomst av crossover ha varit. Man kan nämligen förmoda att det finns en överlappning mellan effekterna av denna kvalitetsaspekt och att död använts som endpoint.

Crossover betyder att man i vissa studier brutit den randomiserade behandlingen och givit den andra vätskan i stället. Läkarkårens tro på albumin som en välgörande substans har tydligen medfört att de svårast sjuka patienterna erhållit albumin trots att de randomiserats till kristalloid vätska. Författarna fann inte något enda fall av crossover där en patient randomiserad till albumin senare fick kristalloid vätska. En sådan brytning av behandlingen är förstås klandervärd om syftet är att belysa mortaliteten, men crossover har troligen förekommit främst i studier som genomförts i andra syften. Studier helt fria från crossover uppvisade därför ingen ökning av mortaliteten (RR 1,04) till skillnad från övriga (RR 1,43).

Risken med albumin var högre i små studier än i »stora«, vilket här innebär fler än 100 patienter. Denna bias kan troligen förklaras av selektiv publicering. Små studier med överraskande resultat blir sannolikt oftare publicerade än studier utan dramatiska moment. Då ingen förväntar sig att albumin är farligt kommer färre små studier med ett neutralt resultat att publiceras, medan artiklar som påvisar en riskökning är mer »säljbara«. Stora studier är mindre känsliga för denna typ av selektion.

Tveksamheter finns

Ingen studie uppfyllde samtliga fyra kriterier på god kvalitet, men författarna laborerar med att beräkna risken i studier som innehåller olika kombinationer av kriterier. Som exempel kan nämnas att de 13 studier där död var slutmättningspunkt och ingen crossover skett visade en relativ risk för albumin på bara 0,90 (95 procents konfidensintervall 0,74–1,09). Läsaren frestas konkludera att den överdödlighet efter albumininfusion som framkommit i tidigare litteraturöversikter tycks vara en följd av att vissa av de ingående studierna haft en mindre lyckad konstruktion.

Ingen betvivlar att metaanalyserna av albumin och kolloiders farlighet bygger på skickligt utförd statistik. Kritiken handlar mer om sunt förnuft i tolkningen av resultaten. Frågor som blindning och crossover är därför inte de enda metodologiska problem som kan skönjas. Flera studier ger doser av albumin som är för stora för att spegla modern praxis, vilket jag tidigare berört [1]. Sådana studier har tyvärr inte exkluderats från denna metaanalys. Mängden infunderat albumin varierar stort mellan studierna, liksom indikationen för terapi samt patienternas allmäntillstånd och grundsjukdomar.

En olycklig egenskap hos den aktuella metaanalysen är också att kontrollbehandlingen kunnat bestå av både isoton och hyperton kristalloid vätska. Koksalt 7,5 procent har prövats i många olika sammanhang under de senaste 20 åren. Mycket talar för att denna hyperton vätska utövar specifika effekter

på såväl perifera ödem som på immunsystem och inflammatoriska kaskader. Att värdera riskerna med 7,5-procentig koksaltlösning tillsammans med isoton koksaltlösning och Ring-erlösning tyder på ett onyanserat synsätt.

Ytterligare en tveksamhet inför denna metaanalys uppkommer då det visar sig att behandlingen i kontrollgruppen inte kunnat bestå av syntetiska kolloider av typ dextran och hydroxyetylstärkelse. Det innebär att eventuella skillnader mellan studiegrupp och kontrollgrupp inte behöver spegla en farlighet hos just albumin, utan lika gärna kan avse kolloida vätskors gemensamma egenskaper, såsom stor volymexpansion per given mängd vätska samt långsam utsöndring. Men samtidigt kan båda grupperna ha erhållit sådana kolloider och dessutom okända mängder erythrocyter, vilket förutom immunologiska effekter också ökar plasmavolymen på ett liknande sätt som kolloida vätskor.

Sammanfattningsvis vet vi att albumin är betydligt dyrare men knappast bättre än syntetiska kolloider. Albumin kan fortfarande vara farligare än kristalloid vätska. Den aktuella metaanalysen påvisar en icke statistiskt säkerställd ökning av dödligheten efter infusion av albumin. Dess värde ligger dock främst i att kvaliteten i ingående studier blottläggs som avgörande för tolkningen av litteraturen.

De två författarnas bakgrund förtjänar en kommentar. De – Mahlou Wilkes och Roberta Navickis – arbetar sedan 20 år tillbaka på ett eget konsultföretag i Kalifornien, Hygiea Associates, som på uppdrag från främst läkemedelsföretag skriver kvalificerade artiklar för biomedicinska tidskrifter. Ingen av dem är läkare. Den föreliggande studien har bekostats av the Plasma Protein Therapeutics Association, vilket är en sammanslutning av företag som arbetar med att utveckla och försälja plasmaprodukter. Även om man kan ha undringar om partiskhet till förmån för albumin, tror jag att vetenskapliga artiklar som emanerar från liknande icke-akademiska sammanhang kommer att bli vanligare i framtiden.

Referenser

1. Hahn RG. Albumin tvivelaktigt som plasmaexpander. *Läkartidningen* 2001;98:806-8.
2. Cochrane Injuries Group Albumin Reviewers. Human albumin administration in critically ill patients: systematic review of randomised trials. *BMJ* 1998;317:235-40.
3. The Albumin Reviewers (Alderson P, Bunn F, Lefebvre C, Li Wan Po A, Lil L, et al). Human albumin solution for resuscitation and volume expansion in critically ill patients (Cochrane Review). In: *Cochrane Library*, issue 1, 2000. Oxford: Update Software.
4. Roberts I, Edwards P, McLelland B. More on albumin. Use of human albumin in UK fell substantially when systematic review was published [letter]. *BMJ* 1999;318:1214-5.
5. Schierhout G, Roberts I. Fluid resuscitation with colloid or crystalloid solutions in critically ill patients: a systematic review of randomised trials. *BMJ* 1998;316:961-4.
6. Choi PT, Yip G, Quinonez LG, Cook DJ. Crystalloids vs colloids in fluid resuscitation: a systematic review. *Crit Care Med* 1999;27:200-10.
7. Wilkes MM, Navickis RJ. Patient survival after human albumin administration. A meta-analysis of randomized, controlled trials. *Ann Intern Med* 2001;135:149-64.

Carl Lindgren, avdelningsläkare, med dr, neuropediatrika programmet, Astrid Lindgrens barnsjukhus, Karolinska sjukhuset, Stockholm (carl.lindgren@ks.se)

Fototerapi eller ej?

Fortfarande saknas konsensus om handläggning av neonatal non-hemolytisk hyperbilirubinemi

Flertalet barnläkare har säkert någon gång under sin karriär ställt sig frågan om beslutet att ljusbehandla ett fullgånget till synes friskt barn som vid blodprovskontroll av bilirubin visat sig överskrida den fastställda behandlingsgränsen verkligen har förhindrat ett fall av framtida neurologisk och/eller kognitiv utvecklingsavvikelse eller bara förorsakat den nyförlösta modern och hennes familj oro och ängslan. Klinisk icterus (serumbilirubin >85 mmol/l) föreligger hos ca 65 procent av alla nyfödda, och upp till 5 procent av skandinaviska nyfödda barn behandlas med någon form av fototerapi. Avsevärt högre siffror anges dock i vissa internationella studier.

I detta nummer av Läkartidningen presenterar Thor Willy Ruud Hansen den nyligen publicerade Cochranerapporten angående utvärdering av fiberoptisk ljusbehandling vid neonatal hyperbilirubinemi. Det är en välkommen utvärdering. Få torde ha bättre förmåga att bedöma denna metaanalys än Ruud Hansen, som har mångårig klinisk och experimentell erfarenhet av bilirubintoxicitet. Med data från 24 randomiserade, kontrollerade studier jämförs effekten vad gäller sänkning av serumbilirubin vid användande av konventionell ljusbehandling (ljusrörslamp ovanför sängen) med den som erhålls vid behandling med de fiberoptiska ljuskällorna Bili-Blanket respektive Wallaby. Även duration av behandling, biverkningar, effekt på anknytningsprocessen mellan mor och barn och kostnadseffektivitet har analyserats.

Cochranerapporten konkluderar att fiberoptiska ljuskällor är likvärdiga med konventionell ljusbehandling vid behandling av preterma barn, men mindre effektiva vid behandling av fullgångna barn. En förklaring till denna diskrepans kan vara att ett proportionerligt mindre hudparti bestrålas med fiberoptiska ljuskällor vid behandling av större barn. Den fiberoptiska apparaturen bestrålar som regel endast rygg respektive mage/bröstkorg men inte extremiteter eller ansikte. Detta minskar teoretiskt den spektrala styrkan (irridans [$\mu\text{W}/\text{cm}^2/\text{nm}$] \times bestrålad hudyta) vid behandling av större barn jämfört med konventionell ljusbehandling.

En fototerapiutrustning som inte berörs i Cochranerapporten, men som på senare år vunnit popularitet i Sverige, särskilt på BB-avdelningar, är BiliBed (Medela Medical, Baar, Schweiz) (Figur 1). Enligt uppgift från fabrikanterna finns idag 60 stycken sådana i kliniskt bruk i landet [PO Schultz, pers medd, 2001]. Trots den flitiga användningen av BiliBed är

Evidensbaserad medicin

Se även artikeln på sidan 4964 i detta nummer.

värdet av denna behandlingsutrustning beskrivet endast i en mindre studie, varvid den i fråga om sänkning av serumbilirubin befanns vara jämförbar med konventionell ljusbehandling för barn med födelsevikt >2 300 gram [1].

Osäkerhet om koncentrationer

Rapporten behandlar inte indikationer för behandling, dvs vilken koncentration av serumbilirubin i förhållande till barnets gestationsålder och postnatala ålder som motiverar behandling. Denna fråga har undersökts i två äldre studier [2, 3], och sannolikheten att idag kunna göra en klinisk studie som på ett avgörande sätt skulle kunna förändra indikationerna för behandling är av naturliga skäl begränsade. Successivt har dock en mer permissiv inställning till non-hemolytisk hyperbilirubinemi vunnit inesteg som ett resultat av de artiklar som kommit från Newmans och Maisels' grupp [4]. Detta har dock fått till följd att det idag råder en osäkerhet om vilka koncentrationer av serumbilirubin som verkligen motiverar behandling. Kärnikterus är idag ytterst ovanligt, och data angående vilka koncentrationer av serumbilirubin i nyföddhetsperioden som kan påverka kognitiva funktioner i senare ålder är motstridiga.

Till följd av detta florerar idag på de svenska barnkliniker-na ett antal olika diagram med gränsvärden för behandling. Dessa diagram relaterar i huvudsak till Dodds översiktsartikel med förslag på diagram [5], som är en vidareutveckling av Hillington Hospital Chart [6]. Detta diagram har i sin tur modifierats efter lokala traditioner, särskilt vad gäller perioden fram till 120 timmars ålder för preterma barn. Resultatet av detta blir att barn som behandlas vid en klinik lämnas obehandlade vid en annan.

Behandlingsgränser

Vid flertalet kliniker torde en gräns på 350 $\mu\text{mol/l}$ för fullgångna barn vid tre dygns ålder resultera i behandling. Vissa



Figur 1. Fototerapi med användande av BiliBed.

kliniker har en gräns på 300 $\mu\text{mol/l}$ vid samtidig förekomst av någon »riskfaktor«, som t ex status post asfyxi eller sepsis. Gränsen på 350 $\mu\text{mol/l}$ motsvarar i runda tal de 20 mg/dl som rekommenderades efter pionjärbetena av Killander och Wishingrad [2, 3].

Vad gäller preterm barn är behandlingsgränserna i huvudsak relaterade till barnets födelsevikt och postnatale ålder. Kurvorna särskiljer vanligtvis inte mellan olika grader av prematuritet. Denna ansats är inte självklart logisk, då den inte tar hänsyn till barnets biologiska mognad, även om kroppsvikt oftast är relaterad till postkonceptionell ålder. Penetrationen av bilirubin till centrala nervsystemet över blod–hjärnbarriären är inte med nödvändighet densamma för ett barn fött efter 32 graviditetsveckor med födelsevikt 1 700 gram som för ett gravt tillväxthämmat barn med samma födelsevikt i graviditetsvecka 36. Det bilirubinvärde som erhålls från laboratoriet kan till följd av kalibreringsteknik och drift i apparaturen ha en viss felprocent på upp till 5 procent i det intervall som är aktuellt under nyföddhetsperioden. Dessutom finns som vid all annan provtagning vissa felkällor vad gäller provtagningsprocedur, transport och hantering av det tagna blodprovet.

Upprepad provtagning för vissa barn

En grupp barn som ofta blir föremål för upprepade provtagning och behandling är de som är födda efter 35–37 graviditetsveckor. Dessa barn vårdas idag som regel på BB-avdelningar där utvecklingen de senaste åren även avseende dessa barn gått mot allt kortare vårdtider. En del av dessa barn, särskilt de som är förlösta med kejsarsnitt, har ibland svårt att etablera en fungerande amning under de första dagarna. Detta kan leda till ett relativt markerat viktfall och en proportionerligt ökad enterohepatisk recirkulation av bilirubin, vilket kan resultera i utveckling av behandlingskrävande hyperbilirubinemi. Amningsobservation, viktkontroll och extra tillmatning vid ett viktfall >7–10 procent av födelsevikten, även vid tillfredsställande blodsockerkoncentration, kan vara ett sätt att förhindra att detta tillstånd uppträder. Tillmatning som kompletterat till amning är idag också en vanlig behandling vid fototerapi.

I den kliniska vardagen påverkas sannolikt vårt ställningstagande till behandling även av hänsyn till andra omständig-

heter än de som är strikt medicinskt relevanta. En sådan faktor kan vara en effekt av den allt kortare vårdtiden efter förlösningen. Ett barn som ligger just under gränsen för behandling på eftermiddagen under sitt andra levnadsdygn kan erhalla »profylaktisk« fototerapi under natten för att undvika att värdet nästa morgon omöjliggör den emotsedda utskrivningen från sjukhuset.

Källa till stor oro och ängslan för familjen

Även om hyperbilirubinemi från klinisk synpunkt kan betraktas som ett relativt harmlöst tillstånd är det en källa till stor oro och ängslan för den nyförlösta modern och hennes familj. Särskilt vid användande av konventionell ljusbehandling gör behandlingen att mor och barn temporärt separeras, vilket kan inverka menligt på såväl anknytningsprocessen som etablering och duration av amningen [7].

Något förvånande är därför resultaten i Cochranerapporten att fiberoptisk fototerapi inte upplevs av föräldrarna som mindre invasiv eller mindre störande för anknytningen än konventionell behandling. Just detta, att behandlingen skulle upplevas mindre traumatisk och kunna utföras i det rum där modern vistades, var ju ett av huvudargumenten vid införandet av fiberoptiska ljuskällor. Detta aktualiserar åter vikten av adekvat utvärdering såväl i medicinskt hänseende som vad gäller emotionella aspekter och kostnad–nyttaanalyser innan nya behandlingsformer, rutiner eller läkemedel införs i kliniskt bruk.

Det är olyckligt att vi idag inte har en rikstäckande konsensus avseende indikationer för fototerapi. Ej heller råder någon bred enighet om vilken apparatur som skall användas, eller om fototerapi skall utföras kontinuerligt eller intermitterant. Cochranerapporten aktualiserar vikten av adekvata utvärderingar av medicinsk apparatur. I en vidare bemärkelse pekar den på vikten av en övergripande översyn av en behandling som i huvudsak grundar sig på ett fixerat – men inte helt säkert exakt – laboratorievärde, speglade en serumkoncentration – men inte vävnadskoncentration – av en biologisk substans vars effekter på den växande hjärnan fortfarande inte är fullständigt klarlagda.

Referenser

1. Lindgren C, Nilson T. Behandling av neonatal hyperbilirubinemi med BiliBed. *Tidsskr Nor Laegeforen* 1999;119(27):4027-9.
2. Killander A, Michaelson M, Muller-Eberhardt U, Sjölin S. Hyperbilirubinemia in full term newborn infants: a follow-up study. *Acta Paediatr Scand* 1963;42:481-4.
3. Wishingrad L, Cornblath M, Takakuwa P, Rosenfeldt IM, Elegant L, Kaufman A, et al. Studies in non-hemolytic hyperbilirubinemia in premature infants: prospective randomized selection for exchange transfusion with observation of the levels of serum bilirubin with and without exchange transfusion and neurological evaluations one year after birth. *Pediatrics* 1965;36:162-72.
4. Newman TB, Maisels MJ. Evaluation and treatment in the term newborn. A kinder, gentler approach. *Pediatrics* 1992;89:809-18.
5. Dodd KL. Neonatal jaundice – a lighter touch. *Arch Dis Child* 1993;68:529-33.
6. Finlay HVL, Tucker SM. Neonatal plasma bilirubin chart. *Arch Dis Child* 1978;53:90-1.
7. Elander G, Lindberg T. Hospital routines in infants with hyperbilirubinemia influence the duration of breastfeeding. *Acta Paediatr Scand* 1986;75:708-12.

Jens Schollin, docent, överläkare, barn- och ungdomskliniken, Regionsjukhuset, Örebro

Cochrane-rapport:

Sackaros minskar smärta hos nyfödda, optimala doser dock okända

■ Under nyföddhetsperioden tas blodprov på alla barn, det så kallade PKU-provet (vanligast vid 3–5 dygns ålder), dessutom ges alla barn K-vitamin vid födelsen. Inte sällan tas flera prov under denna period, och beräkningar från Storbritannien visar att på 15 procent av alla nyfödda tas mer än ett blodprov under första levnadsveckan [1].

Den alltmer utvecklade neonatalvården har också medfört att allt fler prov tas på de barn som föds underburna eller drabbas av neonatal sjukdom. Blodprov och andra prov eller procedurer (intubation, pleuradränage, likvorpunktion) åsamkar det nyfödda barnet smärta, vilken naturligtvis i sig ska behandlas på bästa möjliga sätt. Ett svenskt dokument för behandlingsrekommendationer av smärta hos nyfödda utarbetas av Svensk Barnsmärtförening och beräknas vara klart inom kort.

Historik

Smärta är ett komplext fenomen som oftast uttrycks subjektivt och som är svårt att mäta kvantitativt och kvalitativt. Att uppleva smärta är ett primitivt symptom viktigt för fortsatt överlevnad. Den tidigare uppfattningen att nyfödda inte kan känna smärta har helt reviderats. Beskrivningar finns av hur sprithaltiga drycker, socker, vatten och annat har använts som smärtlindring på nyfödda i anslutning till operation. Neuroanatomiska studier har visat att både centrala och perifera strukturer som är nödvändiga för perception av smärta finns tidigt i graviditeten [2]. Dessutom har det nyfödda barnet sannolikt en ökad smärtekänslighet då inhiberande mekanismer inte är utvecklade. Dessa kunskaper har lett till en förändring i klinisk praxis vid behandling av smärta hos nyfödda [3]. Svårigheterna är dock extra stora vad gäller att mäta och skatta smärta hos nyfödda då de inte kan verbalisera vad de upplever. Reaktion på smärta hos nyfödda kan mätas på tre olika sätt:

1. Beteendemässigt, till exempel skrik eller gråt.
2. Fysiologiska förändringar, till exempel hjärtfrekvens, blodtryck och blodgasmätning.
3. Som förändringar i stresshormoner, som katekolaminer och kortisol.

Fysiologiska indikatorer på smärta, som hjärtfrekvens, räcker inte som enda mätmetod för smärtskattning eftersom dessa oftare är associerade till stress än till smärta. Olika beteendeskolor som inkluderar gråt, skriketid och kroppsrörelser har alltmer kommit att användas. Svårigheterna med dessa är

SAMMANFATTAT

Smärta kan upplevas av alla nyfödda, även mycket för tidigt födda barn.

Bedömning av smärta hos nyfödda bör ske med flerdimensionella skalor där både beteende och fysiologiska parametrar ingår.

Söta lösningar (sackaros) reducerar både fysiologiska parametrar (som hjärtfrekvens) och beteendeparametrar (som gråt och ansiktsuttryck) vid smärtsamma åtgärder hos nyfödda.

Både för tidigt födda barn och barn födda i normal tid har nytta av sackaros som smärtlindring.

Den optimala dosen av sackaros är inte känd.

Fortsatt forskning för att utvärdera effekt, säkerhet och möjlighet till användande av sackaros flera gånger i tillräckligt stora undersökningsmaterial är önskvärd.

Evidensbaserad medicin

att beteenden ändras med gestationsålder och ålder efter födelse, och också kan vara associerat med tidigare smärterfarenhet [4]. Det finns av dessa anledningar för närvarande inget enskilt bra test eller prov på hur smärta ska skattas i nyföddhetsperioden, och ännu mer än för andra åldersgrupper är det därför nödvändigt att använda ett flerdimensionellt sätt att bedöma smärteaktioner hos det nyfödda barnet.

Olika skalor som utvecklats är till exempel »neonatal facial coding system« (NFCS), »children hospital eastern Ontario pain scale« (CHEOPS), »neonatal infant pain scale« (NIPS) och »premature infant pain profile« (PIPP). Den senare, PIPP, används i allt större utsträckning och är validerad för både fullgångna och för tidigt födda barn.

Vid behandling av smärta hos nyfödda är det viktigt att också betona svårigheten att skatta långvarig kronisk smärta,

som vid långvarig respiratorbehandling. De skalor som nu används har i första hand utnyttjats vid så kallade procedur-utlöst smärta, det vill säga smärta orsakad av någon akut intervention, som hälstick eller venpunktion.

Farmakologisk behandling av smärta har getts under längre tid, i första hand med opioider eller paracetamolpreparat. Risken med farmakologisk behandling är biverkningar, vilket lett till att icke-farmakologisk behandling har blivit allt vanligare. Sådan behandling inkluderar bland annat söta orala lösningar, icke-nutritivt sugande och vaggande av barnet. En annan orsak är att det inte är rimligt att ge farmakologisk behandling vid PKU-prov och liknande till fullgångna barn. Andra sätt att minimera smärta är att utforma interventioner på det sätt som minst orsakar smärtreaktion, till exempel att ta blodprov venöst i stället för kapillärt.

Målsättningen med denna översikt är att redogöra för och kommentera en systematisk litteraturanlys från Cochrane-samarbetet avseende användandet av sackaros för analgesi av procedurrelaterad smärta i nyföddhetsperioden [5]. Procedurerna som undersöks är provtagning med hälstick eller venpunktion samt smärta vid intramuskulär injektion. Författare till rapporten är Bonnie Stevens och Arne Ohlsson. Arne Ohlsson är svensk neonatolog nu verksam i Toronto, Kanada.

Rapportens material och metoder

Rapporten utgår från en systematisk genomgång av befintlig litteratur samt sökning i Medline (från 1966 till november 1997), Embase (1993–1997) och Cochrane Controlled Clinical Trials Register. Nyckelord och MeSH-termer vid sökning inkluderade »infant/newborn«, »pain«, »analgesia« och »sucrose«. Personliga publikationer, bibliografier, de mest relevanta tidskrifterna på neonatal- respektive smärtområdet samt konferensrapporter genomsköts manuellt. Urvalet utgjordes av randomiserade kliniska studier där fullgångna eller för tidigt födda barn var provtagna med kapillärprovtagning eller venpunktion eller där intramuskulär injektion (vaccination) gjorts och där sackaros eller vatten/placebo samtidigt givits.

Avsikten var att bestämma effekten av given dos, administrationsätt och säkerhet vad gäller sackaros som smärtlindring till nyfödda. Effekterna mättes som beteendemässiga (duration av skrik, skriketid, ansiktsuttryck) och/eller fysiologiska (hjärtfrekvens, andningsfrekvens, transkutant pO₂, transkutant pCO₂), smärtindikatorer och/eller sammansatta smärtskalor. För att beskriva den metodologiska kvaliteten i varje studie användes standardmetoder från »Neonatal Collaborative Reference Group« som bestod av fyra frågor beträffande varje studies kvalitet avseende:

1. Randomiseringen okänd för försöksledaren och patienten, »blindad randomisering«.
2. Interventionen okänd för försöksledaren och patienten, »blindad intervention«.
3. Fullständigheten vid uppföljning.
4. Utfallet granskat av bedömare där behandling/intervention är okänd för denne, »blindat utfall«.

För varje fråga fanns tre möjliga svar: »går ej att besvara«, »ja« och »nej«.

Den metodologiska kvaliteten bedömdes oberoende av två personer, och vid olika bedömningar togs beslut genom konsensusdiskussion. Varje person extraherade självständigt data från varje artikel. Avsaknad av publicerade medelvärden och standarddeviationer vad gäller utfallet i flera studier förhindrade att metanalys kunde genomföras.

Resultat

Femton studier identifierades för att kunna inkluderas i litteraturgenomgången. Fem av dessa exkluderades. Hos tre av de

exkluderade studierna gavs inte tillräcklig information om hur många nyfödda som randomiserats till studiegrupp respektive kontroll, en studie var inte randomiserad och antalet inkluderade var oklart, i en studie utfördes ingen smärtreducerande procedur. De återstående inkluderade tio studierna beskrivs.

Även om alla studier som inkluderades var randomiserade var informationen om säkerheten i blindningen otillräcklig. Alla studier utom två var dubbelblinda. I dessa två studier förhindrades dubbelblindningen av att flera interventioner gjordes. Få studier rapporterade någon definition av smärta och hur smärtan var relaterad till utfallet. Huvuddelen av studierna gällde hjälpfunktion som smärtprocedur. Det gavs litet information om själva hälstickets utförande, vilket medför svårigheter att jämföra smärtstimulis duration eller frekvens. Observationstidens längd var inte alltid rapporterad, vilket kan ha betydelse för biverkningar. Administrationssättet för sackaros var olika (spruta, droppar eller via napp). Statistisk beskrivning beträffande SD, standardfel eller median rapporterades olika, och ibland bara med grafer utan tillgång till exakta data. Allt detta gjorde att studierna inte kunde kombi-neras, de sammanfattas därför enligt:

Abad 1993 [6]: Medelvärde för procent barn som grät tre minuter efter hälstick reducerades signifikant för sackarosgruppen (0,24 g och 0,48 g) jämfört med vatten. Reduktionen var större för gruppen som fick den högre halten socker.

Abad 1996 [7]: Gråttiden tre minuter efter venpunktion var signifikant lägre om 0,48 g sackaros gavs, men inte för 0,24 g, jämfört med om vatten gavs. För hjärtfrekvens gällde att en signifikant skillnad uppstod om man jämförde sackarosgrupperna tillsammans med gruppen som enbart fick vatten.

Buchner 1995 [3]: Medianökningen i hjärtfrekvens efter hjälpfunktion reducerades signifikant vid 1,0 g sackaros jämfört med vatten, liksom gråttiden.

Gormally 1996 [8]: Procentuellt antal nyfödda som grät vid två respektive tre minuter efter hälstick var signifikant lägre för gruppen som fick sackaros plus bars omkring jämfört med den som fick enbart vatten. Gruppen fick dock inte bättre resultat än grupperna som enbart bars omkring eller gavs sackaros.

Haouari 1995 [1]: Medianen för duration av total gråttid plus första skrikets längd var signifikant lägre vid 1,0 g sackaros liksom medianen för skrik vid en minut respektive två minuter jämfört med om man gav vatten. Vid tre minuter förelåg en signifikant lägre hjärtfrekvens hos dem som fick 1,0 g sackaros.

Johnston 1997 [9]: Vid 30 sekunder efter hjälpfunktion gav 0,01 g sackaros med/utan vaggande en signifikant lägre »ansiktsaktivitet« hos barnet än hos dem som enbart fick vatten.

Ramenghi 1996a [10]: Under fem minuter förelåg en signifikant lägre procentuell skriketid i gruppen som fick 0,25 g sackaros jämfört med gruppen som fick vatten. Samma gällde för duration av första skrik. Sugintensiteten bedömdes vara signifikant lägre i sackarosgruppen jämfört med vattengruppen.

Ramenghi 1996b [11]: Vid tre minuter fanns en signifikant minskning i skriketid och procent nyfödda som skrek efter hjälpfunktion hos dem som fick 0,5 g/1,0 g sackaros och en konventionell söt lösning (Calpol) jämfört med dem som bara fick vatten. Hjärtfrekvensen var signifikant högre efter tre mi-

nuter hos barn som fick vatten jämfört med dem som fick den konventionella lösningen och 1,0 g sackaros. Ett sammansatt smärtindex var högre i vattengruppen jämfört med övriga grupper.

Rushforth 1993 [12]: Inga skillnader kunde noteras mellan 0,15 g sackaros och vatten.

Stevens 1997b [13]: Signifikant lägre PIPP-poäng vid hälstick hos nyfödda som sög på napp doppad i sackaroslösning jämfört med ingen intervention alls.

Rapportens diskussioner och konklusioner

I de bedömda studierna reducerades univariata fysiologiska parametrar (hjärtfrekvens) och beteendeparametrar (gråttid, ansiktsuttryck) liksom multivariata skalor (PIPP) för smärt-skattning av sackaros till nyfödda när hälstick eller venpunktion gjordes. Det finns ingen samstämmighet beträffande vilken dos som är effektiv. I de senare studierna [9, 13] reducerade små volymer sackaros smärta signifikant hos underburna barn. Den metaanalys som gjordes av Stevens [14] visade att 0,18 g sackaros inte hade bättre effekt än vatten. Doser på 0,24 g eller mer var effektiva, men effekten tycktes inte bli bättre med doser över 0,5 g.

Sackaros associeras också med lägre hjärtfrekvens, minskat ansiktsuttryck och mindre motorisk aktivitet efter administration. Den bästa analgetiska effekten tycks erhållas då sackaros ges cirka två minuter före smärtproceduren. Detta tidsintervall kan ge möjlighet för opioidfrisättning.

Inga ogynnsamma bieffekter av att ge sackaros till nyfödda kunde ses i någon studie. Det är emellertid oklart hur lång tid man undersökte barnet avseende eventuella biverkningar. Risken för nekrotiserande enterokolit hos mycket underburna barn har rapporterats från en grupp där upprepade små mängder 20-procentig sackaros blandat med kalciumlaktat gavs [15]. Orsakerna antas vara hyperosmolalitet orsakad av kalciumlaktatet. Av den anledningen bör smärtlindring med söta lösningar studeras särskilt noga för de underburna barnen.

Fem av de inkluderade studierna var utförda på fullgångna barn och fem på underburna. Alla barn var friska och stabila och få var under 37 graviditetsveckor. Underburna barn reagerade på samma sätt som fullgångna även om signalerna ofta var mer subtila och svårfångade. Stevens [16] rapporterade att underburna barn skrek upp till 50 procent mindre än fullgångna barn vid smärtsamma procedurer. Någon kunskap om vilken dos sackaros som är lämpligast för sjuka nyfödda och underburna ger inte de bedömda studierna. Styrkan i studierna i rapporten finns i deras design, då de är prospektiva, dubbelblindade och randomiserade med tydliga kontroll- och interventionsgrupper. En svaghet är dock att flera studier inte kunde användas vid bedömningen. Bland de studier som används finns dock både begränsningar vad gäller begrepp och metodologi, till exempel vad gäller definition av smärta, och hur detta relateras till utfall. Här används skrik eller andel barn som skrek som den vanligaste smärtindikatorn, och få undersökare använde flerdimensionella skalor för att bedöma smärta. Typ av skrik eller gråt har använts för analys vad gäller till exempel intensitet, melodi och reaktion på smärta, men detta användes inte här. Skrik och gråt är också svårt att använda för sjuka nyfödda, till exempel barn som respiratorvårdas. Därför föreslår nyare forskning flerdimensionell bedömning av smärta i nyföddhetsperioden [4].

Olika metoder att utlösa smärta har studerats utan att information i alla studier har givits om hur denna procedur utförts. Detta gör det också svårt att jämföra de olika smärtutlösande procedurerna. Hur barnen förbereds inför smärtproceduren har också betydelse för hur smärtreaktionen utfaller.

Av betydelse är också att studierna har olika sätt att förse barnen med sackaros. Om det ges med napp kan denna i sig också bidra till smärtlindring via icke-nutritivt sugande [17]. Flera undersökare har funnit att napp minskar hjärtfrekvens, metabolism och höjer smärtröskeln. Fysisk kontakt med modern har dock inte visat sig ha effekt på kortisol svar, hjärtfrekvens och syrgassaturation [18]. Peroralt sackaros har till skillnad från kontakt en bestående lugnande effekt även efter den specifika smärtstillande proceduren.

Som konklusion sammanfattar författarna att sackaros minskar det nyfödda barnets smärta vid de beskrivna procedurerna. Mycket små doser rapporteras vara effektiva hos för tidigt födda barn, medan större dos behövs för fullgångna barn. En optimal dos har inte kunnat identifieras och ingen säker viktig biverkan har hittills rapporterats. Fortsatta studier bör fokuseras på effektivitet och säkerhet samt beskriva smärtstimuli och interventionen mer noggrant. Tillräckligt stora material samt användandet av flerdimensionell utvärdering av smärta med validerade variabler är viktiga liksom utvärdering av repeterade doser av sackaros. Användandet av sackaros hos underburna och sjuka nyfödda behöver också ytterligare belysas. Högsta möjliga metodologiska kvalitet med identiska utfallsvariabler är viktigt för att kunna genomföra framtida metaanalyser.

Egna kommentarer

Idag tvekar ingen om att nyfödda barn, inklusive underburna, känner smärta. Att behandla kronisk och akut smärta på ett effektivt sätt är ett gemensamt mål som många forskargrupper arbetar för. Vi vet att cirka 50 procent av svenska nyföddhetsavdelningar skattar och behandlar akut smärta på ett organiserat och strukturerat sätt. Detta är en alldeles för låg siffra, den borde vara 100 procent. För kronisk smärta är motsvarande siffra okänd.

Svårigheterna är många vad gäller skattning och behandling av smärta hos nyfödda, vilket belyses väl i den refererade undersökningen. För det första gäller hur smärta skall skattas. Det är uppenbart att enbart fysiologiska parametrar inte räcker, utan smärta måste bedömas på ett flerdimensionellt sätt där beteendeparametrar och fysiologiska parametrar inkluderas. Flera sådana skalor har konstruerats, och den mest använda idag är den så kallade PIPP-skalan. Denna skala används idag också för upprepade bedömningar vid kronisk smärta hos nyfödda barn.

Ytterligare studier behövs vad gäller hur smärta ska bedömas, särskilt gäller detta för kronisk smärta. En ny gammal metod som idag prövas av några nordiska forskargrupper är mätning av galvanisk hudresistans (GSR), vilket egentligen betyder att man använder en »lögndetektor«. Minskning/ökning i elektrisk ledningsförmåga i huden påverkas av det autonoma nervsystemet genom svettning, vilket i sin tur påverkas av smärtreaktioner. Resultaten hittills är lovande och kan komma att användas även hos underburna barn vid kronisk smärta, men detta behöver utvärderas mer [Hanne Storm, Rikshospitalet Oslo, pers medd, 2001]. Farmakologisk behandling av akut smärta baseras på användandet av opioider och paracetamol. Kronisk smärta behandlas ofta på samma sätt, med regelbundet användande av dessa medel. Kontinuerlig låg dos morfin utvärderas för närvarande i en internationell multicenterstudie, NEOPAIN-studien, där två svenska enheter deltar. De barn som utvärderas är underburna barn som respiratorbehandlas. Det finns klara svårigheter att utvärdera effekten av lågdosmorfin, men variabler som överlevnad, förekomst av hjärnblödning och neurologisk utveckling används.

I en nyligen publicerad studie visar Grunau och medarbetare att användandet av morfin vid smärtsamma procedurer hos underburna gör att dessa uppvisar ett »normalt« smärtbe-

teende vid fyra månaders ålder i motsats till dem som inte fått morfin [19].

Den icke-farmakologiska behandlingen av smärta har inkluderat bland annat söta lösningar, vaggande, icke-nutritivt sugande på napp och amning. Söta lösningar, och då främst sackaros och glukos, får anses säkert visats ha effekt vad gäller reaktioner och beteenden som sätts i samband med smärta hos nyfödda barn, och då särskilt fullgångna. Särskilt är smärta vid hålstick och venpunktion studerad. Mekanismen sker sannolikt via frisättning av opioidliknande substanser, men även via andra mekanismer. Frisättning av tarmhormon, somolecystokin, kan till exempel verka smärtlindrande. En omedelbar effekt kan oftast ses efter att det nyfödda barnet fått en söt lösning, utöver den effekt som ses efter några minuter och som anses vara opioidmedierad. Att bara suga på napp eller finger kan i sig också verka smärtlindrande, kanske enbart genom att man avleder barnet. Den optimala mängden söt lösning är inte känd men däremot, som denna genomgång visar, att den söta lösningen inte får ges i för liten mängd eller vara för svag. Att amma har ansetts, och anses, vara smärtlindrande. Några kontrollerade studier inom detta område finns inte, men flera är på väg [Ricardo Carbajal, Poissy, Frankrike, pers medd, 2001, Maria Gradin, Örebro, pers medd, 2001]. Utöver detta kan smärta hos nyfödda barn förebyggas genom olika andra aktiviteter. Vi vet att venpunktion är mindre smärtsamt än hålstick, men om vi använder söta lösningar blir skillnaden mellan hålstick och venpunktion mindre [20]. Vi vet också att Emlakräm fungerar som smärtlindrande, även på fullgångna barn, men att söta lösningar verkar ha en ännu bättre effekt. En kombination av Emlakräm och söta lösningar kan därför vara ett framtida alternativ.

Allt detta visar att behovet av fortsatt forskning är stort, men de framsteg som gjorts de senaste tio åren vad beträffar insikten om att smärta kan upplevas och behandlas hos nyfödda är betydande.

Referenser

1. Haouari N, Wood C, Griffiths G, Levene M. The analgesic effect of sucrose in full term infants: a randomised controlled trial. *BMJ* 1995;310:1498-500.
2. Anand KJS, Hickey PR. Pain and its effects in the human neonate and fetus. *N Engl J Med* 1987;317:1321-9.
3. Buchner HU, Moser T, von Siebenthal K, Keel M, Wolf M, Duc G. Sucrose reduces pain reaction to heel lancing in preterm infants: A placebo-controlled, randomized and masked study. *Pediatr Res* 1995;38:332-5.
4. Stevens B. Composite measures of pain in children. In: Finley GA, McGrath PJ, editors. *Measurement of pain in infants and children*. Progress in pain research and management. vol. 10. Seattle: IASP Press, 1997. p 161-77.
5. Stevens B, Ohlsson A. Sucrose for analgesia in newborn infants undergoing painful procedures. The Cochrane library: <http://www.update-software.com/clibhome/clib.htm>
6. Abad F, Diaz NM, Domenech E, Robayna M, Rico J, Arrecivita A, et al. Attenuation of pain related behavior in neonates given oral sweet solutions [abstract]. Paris: 7th World congress of pain; 1993.
7. Abad F, Diaz NM, Domenech E, Robayna M, Rico J. Oral sweet solution reduces pain-related behavior in preterm infants. *Acta Paediatr* 1996;85:854-8.
8. Gormally SM, Barr RG, Young SN, Alhawaf R, Wersheim L. Combined sucrose and carrying reduces newborn pain response more than sucrose or carrying alone [abstract]. *Arch Pediatr Adolesc Med* 1996;150:47.
9. Johnston C, Stremmer R, Stevens B, Horton L. Effectiveness of oral sucrose and simulated rocking on pain response in preterm neonates. *Pain* 1977;72:193-9.
10. Remenghi LA, Wood CM, Griffith GC, Levene MI. Reduction of pain response in premature infants using intraoral sucrose. *Arch Dis Child* 1996;74:F126-8.
11. Remenghi L, Griffith G, Wood C, Levene M. Effect on non-sucrose sweet tasting solution on neonatal heel prick responses. *Arch Dis Child* 1996;74:F129-31.
12. Rushforth JA, Levene MI. Effect of sucrose on crying in response to heel stab. *Arch Dis Child* 1993;69:388-9.
13. Stevens B, Johnston C, Franck P. Nonpharmacologic interventions for decreasing procedural pain in preterm neonates [abstract]. Helsinki: Fourth international symposium on pediatric pain; 1997.
14. Stevens B, Taddio A, Ohlsson A, Einarson T. The efficacy of sucrose for relieving procedural pain in neonates – a systematic review and meta-analysis. *Acta Paediatr* 1997;86:837-42.
15. Willis D, Chabot R, Radde I, Chance G. Unsuspected hyperosmolality of oral solutions contributing in necrotizing enterocolitis in very-low-birth-weight infants. *Pediatrics* 1997;60:535-8.
16. Stevens B, Johnston C, Horton L. Factors that influence the behavioral responses of premature infants. *Pain* 1994;59:101-9.
17. Campos RG. Rocking and pacifier; two comforting interventions for heel stick pain. *Res Nurs Health* 1994;17:321-31.
18. Gunnar MR. Reactivity of the hypothalamic-pituitary-adrenocortical system to stressors in normal infants and children. *Pediatrics* 1992;90:491-7.
19. Grunau RE, Oberlander TF, Whitfield MF, Fitzgerald C, Lee SK. Demographic and therapeutic determinants of pain reactivity in very low birth weight neonates at 34 weeks postconceptional age. *Pediatrics* 2001;107:105-12.
20. Eriksson M, Gradin M, Schollin J. Oral glucose and venepuncture reduce blood sampling pain in newborns. *Early Hum Dev* 1999;55:211-8.

SUMMARY

Cochrane review: Sucrose reduces pain in newborns, optimal doses however not known

Jens Schollin

Läkartidningen 2001;98: 4970-5

Management of pain in newborns is less than optimal. We know for certain that newborns, even if born extreme premature, can feel pain. The assessment of pain is difficult and should include both physiological and behavioural parameters. The Cochrane Collaborative presents a review considering the efficacy of sucrose as analgesia in newborns undergoing painful procedures. The review shows that sucrose is efficient for procedural pain reduction even for premature newborns. The optimal doses are however not known and further studies must focus on efficacy and safety.

Correspondence: Jens Schollin, Barn- och ungdomskliniken, Örebro läns landsting, SE-701 85 Örebro, Sweden.

Thor Willy Ruud Hansen, seksjonsoverlege, dr med, Nyfødtsseksjonen, Barneklubben, Rikshospitalet, Oslo, Norge
(t.w.r.hansen@klinmed.uio.no)

Fiberoptisk fototerapi ved neonatal ikterus

Kommentar til en Cochrane-rapport

II Gulsott hos nyfødte er trolig det vanligste tegn som medfører medisinske tiltak i form av prøvetagning og eventuelt behandling i denne aldersgruppen. Ikterus hos nyfødte er hos de fleste et normal-fysiologisk fenomen. Under spesielle omstendigheter, f.eks. ved sterk hemolyse (Rh-immunisering), vil bilirubin kunne trenge inn i hjernen i slik grad at det medfører varig hjerneskade (kjerneikterus). Målet med behandling av gulsott hos nyfødte er å avverge slik skade.

Den terapeutiske effekten av lys ved gulsott hos nyfødte ble oppdaget i England på 1950-tallet [1]. Fototerapi er i dag en av de vanligste behandlinger i nyfødttmedisinen. Konvensjonell lysbehandling innebærer at barnet legges i en seng som belyses av en serie lysstoffrør eller en rekke halogenlamper. Behandlingen vil ofte medføre at barnet fysisk adskilles fra moren. Utviklingen av fiberoptiske behandlingssystemer ble derfor hilst velkommen fordi det førte til at det ble mulig å behandle barnet i samme rom som moren (Figur 1). Man antok at dette ville medføre mindre forstyrrelser av den viktige tilknytningsprosessen mellom mor og barn i de første døgn etter fødselen.

Metaanalysens materiale og metode

Analysen [2] har tatt utgangspunkt i Cochrane-samarbeidets oversikt over kontrollerte studier (Cochrane Controlled Trials Register) samt de elektroniske databasene Medline og Embase for årene 1989–2000. Videre har man funnet artikler fra referanselistene til ovenstående og fra oversiktsartikler. Man har også gått gjennom abstraktene fra møter i The European Society for Paediatric Research og The Society for Pediatric Research. Endelig har man kontaktet forfatterne til de identifiserte artiklene for å få oversikt over eventuelle upubliserte data.

Ved denne fremgangsmåten fant man i alt 31 studier. Av disse ble syv ekskludert fra videre vurdering fordi man ikke kunne være sikker på at inklusjonskravene var tilfredsstillende. Man inkluderte randomiserte og kvasi-randomiserte studier som behandlet fiberoptisk fototerapi av barn opp til 28 dagers alder. Som konvensjonell fototerapi definerte man behandling med lysstoffrør eller rekke av halogenlamper.

De resultatmål man vurderte omfattet: endringer i serum bilirubin per tidsenhet ($\mu\text{mol/l/time}$), varighet av fototerapi

SAMMANFATTAT

24 studier tilfredsstilte inklusjonskriteriene for Cochrane-rapporten om bruk av fiberoptisk fototerapi ved gulsott hos nyfødte.

Fiberoptisk fototerapi har effekt på gulsott, og kan ha en plass i behandlingen av neonatal ikterus.

Hos barn født nær termin er fiberoptisk behandling dårligere enn konvensjonell fototerapi, mens den hos premature barn er jevngod.

Når fiberoptisk behandling kombineres med konvensjonell behandling, oppnås bedre effekt enn med konvensjonell behandling alene.

Fiberoptisk behandling er ikke dokumentert å være mindre forstyrrende enn konvensjonell fototerapi når det gjelder foreldrenes opplevelse av og tilknytning til barnet.

Det er fortsatt stort behov for forskning vedrørende fototerapi av nyfødte med gulsott.

Evidensbasert medisin

Se även medicinsk kommentar i detta nummer.

(timer), forekomst av kjerneikterus (procent), forekomst av behandlingssvikt (procent behov for utskiftingstransfusjon eller tillegg av fototerapienheter), forekomst av bivirkninger, foreldre-barn tilknytning, sykepleiernes tilfredshet med behandlingen, mødrenes tilfredshet med behandlingen, og kostnadseffektivitet. Analysen har skilt mellom barn født nær termin (≥ 37 ukers gestasjonsalder) og premature barn.

Av de 24 studiene som ble analysert var det en studie som



Figur 1. En fiberoptisk fototerapi-enhet (BiliBlanket, Ohmeda Medical) vist i bruk. Som det fremgår kan barnet være hos sin mor under behandlingen.

sammenlignet fiberoptisk fototerapi med ingen behandling, mens 17 sammenlignet fiberoptisk og konvensjonell fototerapi. 7 studier sammenlignet konvensjonell fototerapi med kombinasjonen av konvensjonell pluss fiberoptisk fototerapi, og 1 studie sammenlignet konvensjonell med dobbel fiberoptisk fototerapi. Kun 2 studier sammenlignet ulike fabrikater av fiberoptiske fototerapienheter. Siden studiene tok for seg ulike aspekter, har man i metaanalysen splittet data fra noen enkeltstudier og brukt disse dataene som om de kom fra to separate undersøkelser.

Utgangsmaterialet for analysen var svært heterogent. Studiene inkluderte to ulike fiberoptiske enheter (BiliBlanket og Wallaby) og konvensjonell fototerapi ble bl a gitt med halogenlamper såvel som fluorescerende lamper, og med blått eller hvitt lys såvel som kombinasjoner av disse. Irradiansen fra de to fiberoptiske enhetene er svært forskjellig, i hvert fall slik det oppgis fra fabrikantene, og irradiansen ved konvensjonell fototerapi er også svært variabel. Noen av disse forholdene var ikke i tilstrekkelig grad forutsett da metaanalysen ble planlagt. Derfor har man forsøkt å splitte materialet i subgrupper ved post hoc-analyser.

Resultatene av metaanalysen

Resultatene av metaanalysen presenteres gruppert etter elleve forskjellige sammenligninger. Selv om de 24 inkluderte studiene omfattet i alt 1 753 barn, omfatter enkelte av sammenligningene bare en til to studier, og det kan diskuteres hvorvidt slike sammenligninger i det hele tatt er verdt å presentere i en metaanalyse. Således var det kun én studie som sammenlignet fiberoptisk fototerapi (Wallaby) med en kontrollgruppe uten behandling. Som man med rimelighet kunne forvente var endringen i serum bilirubin over tid større i den behandlede enn i den ikke-behandlede gruppen.

Flertallet av studiene tok for seg sammenligningen mellom fiberoptisk og konvensjonell fototerapi. Reduksjonen i serum bilirubin per tidsenhet var signifikant større i gruppene som fikk konvensjonell fototerapi vs fiberoptisk. Gruppene som fikk fiberoptisk fototerapi hadde signifikant større behov for fototerapi med tilleggsenheter, og man kunne påvise en trend i samme retning når det gjaldt behov for utskiftningstransfusjoner, men denne trenden nådde ikke signifikans. Trans-epi-

dermalt væsketap og blodstrøms hastigheten i a mesenterica var signifikant høyere hos de barna som fikk fiberoptisk behandling.

Premature barn ble studert i en egen analyse. For disse barna tydet analysen på at fiberoptisk fototerapi ga resultater som var sammenlignbare med konvensjonell behandling. Således var det ingen forskjell mellom behandlings-gruppene i forhold til endring i serum bilirubin over tid, varighet av behandling eller behov for tilleggsbehandling.

I 13 studier hadde man sammenlignet BiliBlanket med konvensjonell fototerapi, mens Wallaby var sammenlignet med konvensjonell behandling i seks studier. Dessuten forelå det to studier hvor man hadde sammenlignet BiliBlanket og Wallaby direkte. Metaanalysen ga ikke holdepunkter for signifikante forskjeller i effektivitet mellom de to apparatene.

Dobbelt fiberoptisk fototerapi (bruk av to BiliBlanket) var rapportert i kun én studie, der man ved sammenligning med konvensjonell fototerapi ikke fant signifikante forskjeller mellom gruppene. I seks studier hadde man sammenlignet kombinasjonen av fiberoptisk pluss konvensjonell fototerapi med konvensjonell fototerapi alene. Metaanalysen viste en trend i retning bedre effekt av kombinasjonsbehandlingen, men forskjellene var ikke signifikante.

Metaanalysens diskusjon og konklusjoner

Søkestrategien avdekket et stort antall studier som inkluderte i alt 1 753 barn. Forfatterne mener at det er usannsynlig at de skulle ha oversett studier av betydning. Studiene holdt gjennomgående høy metodologisk kvalitet med tanke på tilfeldig, »random«, allokering av pasientene. Viktige problemer i analysen har vært bruken av forskjellige fiberoptiske apparater, samt at irradiansen fra disse oppgis så forskjellig fra fabrikantene. Videre har det vært stor variasjon i måten den konvensjonelle fototerapien har vært brukt på.

Forfatterne trekker følgende konklusjoner av sin analyse:

1. Fiberoptisk behandling med Wallaby er bedre enn ingen behandling.
2. Fiberoptisk behandling er dårligere enn konvensjonell lysbehandling. Dog er effekten jevn god ved behandling av premature og når man bruker to BiliBlanket samtidig.
3. Kombinasjon av fiberoptisk og konvensjonell fototerapi er mer effektiv i å senke serum bilirubin enn konvensjonell behandling alene.
4. Det er ingen holdepunkter for at det ene fiberoptiske apparatet er bedre enn det andre.
5. Det er ingen holdepunkter fra randomiserte undersøkelser for at fiberoptisk fototerapi er bedre med tanke på barn-foreldre tilknytning eller tilfredshet hos mødre eller sykepleiere.

Forfatterne konkluderer med at fiberoptisk fototerapi har en plass i behandlingen av gulsott hos nyfødte. Behandlingstiden blir trolig lengre enn ved konvensjonell behandling, men dette mener de er uten betydning i de fleste tilfeller. Hos premature barn er fiberoptisk behandling like effektiv som konvensjonell behandling, og de to behandlingsmåtene kan derfor brukes om hverandre. Hos et barn med serum bilirubin i nærheten av utskiftningsgrensen vil en kombinasjon av to fiberoptiske enheter eller en konvensjonell og en fiberoptisk enhet trolig senke serum bilirubin konsentrasjonen raskere enn konvensjonell behandling alene.

Forfatterne tilrår videre forskning på dette feltet. Aktuelle forskningsområder er fiberoptisk behandling ved hemolytisk ikterus, effekter på barn-foreldre tilknytning, »bonding«, til-

fredshet hos mødre og pleierstab samt økonomisk sammenligning av de to behandlingformene.

Min kommentar

Gulsott er en av de vanligste behandlingstrengende tilstander hos nyfødte. Fototerapi er en enkel behandling som likevel legger beslag på betydelige ressurser når man tar i betraktning det relativt store antall barn som behandles. Etter hvert som den fryktede komplikasjonen til gulsott, kjerneikterus, omtrent forsvant i den industrialiserte verden, begynte man i økende grad å stille spørsmål ved indikasjonene for behandlingen [3]. Selv om behandlingen oppfattes som uskadelig, har man likevel bekymret seg for mulige bivirkninger. En av ulempene ved konvensjonell fototerapi var (og er) at det nyfødte barnet oftest må tas fra moren og legges i »lysseng« på et eget rom. Denne adskillelsen av mor og barn kunne tenkes å ha negativ virkning på den viktige tilknytningsprosessen mellom moren og barnet i de første levedøgnene.

Da fiberoptiske fototerapiapparater kom på markedet, var en av de sentrale salgargumentene nettopp at fototerapien nå kunne skje med barnet på mors rom, fordi moren ikke ville få sin nattesøvn forstyrret av det sterke lyset fra en lysseng. Utviklingen av holdningene til fødekvinne ved mange fødeavdelinger, for eksempel uttrykt ved »mor–barn vennlig initiativ«, medførte trolig at et slikt salgargument virket mer overbevisende enn det ville ha vært i en tid da holdningene var mindre sentrerte på mor–barn dyaden.

Mills og Tudehope har påtatt seg en viktig oppgave ved å sammenfatte og metaanalysere de studiene som har vært gjort med tanke på fiberoptisk fototerapi [2]. Utfordringen må ha vært meget stor fordi materialet var meget heterogent, og forfatterne kommenterer dette flere steder. Jeg synes forfatterne er kommet godt fra oppgaven, og de har gjort et meget stort arbeid med å spore opp informasjon i tillegg til den publiserte. Imidlertid har jeg noen kritiske kommentarer.

Hovedproblemet i analysen

Hovedproblemet i analysen har altså vært materialets heterogenitet. For det første er hverken indikasjoner eller prosedyre for fototerapi ved neonatal ikterus standardisert. Det er derfor ikke overraskende at studier viser stor spredning i forhold til praksis [4, 5]. Gitt at spredningen i behandlingsindikasjoner er så stor at et barn med en gitt serum bilirubin konsentrasjon kan bli utsatt for en utskiftningstransfusjon på avdeling »A« mens man ved avdeling »B« ikke en gang ville ha startet fototerapi på dette nivået [4], og at irradiansen ved konvensjonelle fototerapioppsett kan variere med en faktor på opp mot 10 [5], kan man spørre seg om en meningsfylt metaanalyse av behandling i det hele tatt kan gjennomføres.

Derne er det et problem at man ikke har vært i stand til å måle det resultatet som man egentlig er ute etter, nemlig redusert forekomst av kjerneikterus (eller nevrotoksisitet). Kjerneikterus er (og bør være) så sjelden at en studie som målte dette som effekt ikke ville være mulig. Det finnes heller ingen andre alment akseptert mål på bilirubin-indusert nevrotoksisitet. Måling av endringer i serum bilirubin har klare svakheter som surrogat-resultat, men det finnes dessverre ingen andre parametre som er vist å være bedre. Etter at fototerapi som behandlingsform ble gjennomført i den industrialiserte verden, forsvant imidlertid kjerneikterus praktisk talt. Man kan derfor argumentere for at verdien av fototerapi slik sett er dokumentert, og at det er akseptabelt å bruke serum bilirubin verdier som surrogatmål ved sammenligning av ulike former for fototerapi.

Et annet problem er at irradians både måles og rapporteres på ulik måte i litteraturen. Det er derfor svært vanskelig å være sikker på at tall for irradians fra ulike publikasjoner i det hele

tatt er sammenlignbare. Ideelt bør sammenligning av lysdoser fra forskjellige fototerapienheter skje ved kalkulasjon av spektral styrke, »spectral power« [6]. Spektral styrke er produktet av spektral irradians ($\mu\text{W}/\text{cm}^2/\text{nm}$) og størrelsen på det bestrålte hudområdet. Fordi fiberoptiske apparater bestråler et relativt lite område av huden, blir den spektrale styrken dårlig selv om den rapporterte spektrale irradiansen for apparatene ser tilforlataelig ut.

Første del av konklusjon 2 i Cochrane-analysen er således i samsvar med hva man trolig ville ha stillet som hypotese ut fra fysikkens lover. Den neste setningen i konklusjon 2 kan problematiseres. Det er rimelig at effekten av fiberoptisk fototerapi er relativt bedre hos premature enn hos større barn. Det hudområdet som en fiberoptisk matte bestråler vil utgjøre en større andel av deres totalt tilgjengelige hudoverflate enn hos de større barna, og derfor vil også en høyere andel av deres totale bilirubin-pool være tilgjengelig for isomerisering. Om dette blir like godt som konvensjonell fototerapi vil bli avhenge av hvordan sistnevnte er konfigurert, og dette har ikke metaanalysen hatt mulighet til å korrigerer for. Tilsvarende innvending vil gjelde for konklusjonen om at til Bili-Blanket er ekvivalent med konvensjonell fototerapi.

Når det gjelder konklusjon 1, er denne ikke særlig viktig eller interessant for brukeren. Da de fiberoptiske apparatene kom på markedet var konvensjonelle enheter rådende behandlingsstandard. Ved innføring av f eks ny medikamentell behandling er det en regel at den nye behandlingen må vise sin berettigelse i sammenligning med det beste av eksisterende behandling. Med en slik sammenligning ville de fiberoptiske apparatene ha falt igjennom.

Konklusjon 3 er trolig holdbar (enn skjønt kontrasten ikke var statistisk signifikant). Kombinasjonen av fiberoptisk og konvensjonell behandling vil oftest skje ved at barnet legges på en fiberoptisk lysmatte og så får konvensjonell fototerapi ovenfra. Den totale spektrale styrken vil derved økes tilsvarende produktet av spektral irradians og størrelse av det området som lysmatten bestråler. Behandlingseffekten vil øke tilsvarende økningen i spektral styrke. Konklusjonene 4 og 5 tror jeg er holdbare ut fra det som i dag finnes av kunnskap.

Cochrane-analysens forfattere konkluderer med at fiberoptisk fototerapi har en plass i behandlingen av gulsott hos nyfødte, og hevder at den lengre behandlingstiden er uten betydning. Betydningen av behandlingstidens lengde kan utvilsomt diskuteres. Med dagens økte press i retning korte sykehusopphold og rask utskrivelse, er det ikke uten videre gitt at økt oppholdslengde er uproblematisk. Forskjellene i behandlingstid kan være meget store med ulike oppsett. I en nylig publisert studie fra Italia var gjennomsnittlig varighet av fototerapi 80 timer [7]. I vår egen neonatalavdeling er gjennomsnittlig total varighet av fototerapi 20 timer for hvert barn som får slik behandling, og i vår barselavdeling 17 timer [8, 9].

Spørsmålet ikke endelig besvart

Spørsmålet om hvilken plass fiberoptisk fototerapi skal ha i behandlingen av nyfødte med ikterus er ikke endelig besvart, og denne Cochrane-analysen gir heller ikke svaret. Det er klart at konvensjonell fototerapi med et optimalisert oppsett er mer effektiv enn fiberoptisk terapi. Ulempen er at barnet fjernes fra sin mor, at øynene tildekkes, og kanskje kan det være virkninger på andre fysiologiske funksjoner som er uønsket hos spesielle pasienter. Spørsmålet om hvor viktige disse innvendingene er i forhold til den lengre varigheten av behandlingen når denne utføres med fiberoptiske apparater, må den enkelte lege fortsatt besvare skjønnsmessig i forhold til hver enkelt pasient.

Hvorvidt fiberoptisk behandling innebærer noen fordeler i forhold til konvensjonell behandling når det gjelder tilknyt-

ningsprosessen mellom mor og barn, såvel som når det gjelder andre psykologiske og fysiologiske funksjoner, avventer fortsatt et vitenskapelig begrunnet svar. Ut fra dette bør det være klart at lysbehandling av nyfødte fortsatt trenger videre vitenskapelige studier. Det er også grunn til å tro at den praktiske gjennomføringen av lysbehandling kan optimaliseres mange steder [10].

Referanser

1. Cremer RJ, Perryman PW, Richards DH. Influence of light on the hyperbilirubinemia of infants. *Lancet* 1958;1:1094-7.
2. Mills JF, Tudehope D. Fiberoptic phototherapy for neonatal jaundice (Cochrane Review). In: *The Cochrane Library*, issue 1. Oxford: Update Software; 2001.
3. Newman TB, Maisels MJ. Evaluation and treatment of jaundice in the term newborn: a kinder, gentler approach. *Pediatrics* 1992;89:809-18.
4. Hansen TWR. Therapeutic approaches to neonatal jaundice: An international survey. *Clin Pediatr* 1996;35:309-16.
5. Christensen T, Amundsen I, Kinn G, Kjeldstad B. Terapienheter og belsingsforhold ved lysbehandling av nyfødte med hyperbilirubinemi. SIS-rapport 1992:2. Østerås, Norway: National Institute of Radiation Hygiene.
6. Maisels MS. Neonatal hyperbilirubinemia. In: Klaus MH, Fanaroff AA. *Care of the high-risk neonate*. Fifth edition. Philadelphia: WB Saunders Company; 2001. p. 324-62.
7. Pezzati M, Biagiotti R, Vangi V, Lombardi E, Wiechmann L, Rubaltelli FF. Changes in mesenteric blood flow response to feeding: conventional versus fiber-optic phototherapy. *Pediatrics* 2000;105:350-3.
8. Hansen TWR. Phototherapy and mesenteric blood flow – too much, not enough, or the wrong kind? *Pediatrics* 2001;107:610-1.
9. Meslo M, Tveiten L, Hansen TWR. Effektivisering av lysbehandling med enkle midler. *Jordmorbladet* 2000;7:18-21.
10. Hansen TWR. Phototherapy for neonatal jaundice – still in need of fine tuning. *Acta Paediatr* 2000;89:770-2.

SUMMARY

Fiberoptic phototherapy for neonatal jaundice – comments on a Cochrane report

Thor Willy Ruud Hansen

Läkartidningen 2001;98:4964-9

Twenty-four studies filled the inclusion criteria for a Cochrane report on the use of fiberoptic phototherapy for neonatal jaundice. Fiberoptic phototherapy lowers serum bilirubin and may have a place in the treatment of neonatal jaundice. In term and near-term infants fiberoptic phototherapy is inferior to conventional phototherapy, whereas in premature infants the effects are comparable. A better effect is achieved when conventional and fiberoptic phototherapy are combined as compared to conventional phototherapy alone. Fiberoptic phototherapy has not been shown to interfere less with parent-infant bonding than conventional phototherapy. There continues to be a great need for research concerning phototherapy for neonatal jaundice.

Correspondence: Thor Willy Ruud Hansen, Dept of Pediatrics, Rikshospitalet, NO-0027 Oslo, Norway
(t.w.r.hansen@klinmed.uio.no)

Särtryck

Läkartidningen

I samband med 1994 års handikappreform beslutade riksdagen att stimulera utvecklingen av rehabilitering och habilitering med 1,4 miljoner kronor, bidrag som Socialstyrelsen fördelade och följde upp. Ett axplock av de medicinskt fokuserade projekt som bedömts vara av riksintresse har presenterats som separata artiklar i *Läkartidningen* under 1996 och 1997 och i serieform under 1999–2001.

Denna serie har nu samlats i ett särtryck, som av Socialstyrelsens projektgrupp kompletterats med två rapporter som inte redovisats i tidningen.

Särtrycket täcker ett brett spektrum av rehabilitering och habilitering, t ex vid slaganfall, demens och svåra hjärnskador, för barn med särskilda behov och de många små grupperna med ovanliga funktionshinder. Skriften har distribuerats av Socialstyrelsen. En mindre upplaga finns att beställa från *Läkartidningen*.

Priset är 100 kronor.



Habilitering och rehabilitering

Beställer härmed.....ex av
Habilitering och rehabilitering

.....
namn

.....
adress

.....
postnummer

.....
postadress

Insändes till LÄKARTIDNINGEN
Box 5603
114 86 Stockholm

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Baldvin Jónsson, med dr, överläkare, neonatalprogrammet, Astrid Lindgrens barnsjukhus, Karolinska sjukhuset, Stockholm.
(baldvin.jonsson@kbh.ki.se)

Kommentar till en Cochraneöversikt

HFOV eller konventionell ventilation för nyfödda med svår andningssvikt

II Konventionell respiratorbehandling av nyfödda barn använder sig traditionellt av tids- och tryckstyrd respirator som arbetar med frekvenser mellan 30 och 80 andetag/minut. Konventionell respiratorbehandling kan skada lungvävnaden genom uttöjning av alveolarmembranen när höga lufttryck används (volutrauma) [1]. HFOV (högfrekvent oscillatorventilation)-tekniken utvecklades för att minimera vävnadsskada vid mekanisk ventilation, genom att gasutbytet sker utan kraftiga svängningar av luftvägstrycket. Vid HFOV genereras en oscillerande våg med frekvens 5–15 Hz (300–900 svängningar per minut). Ett kontinuerligt luftvägstryck används för att maximera alveolär yta och förbättra syresättningen, medan amplituden i den oscillerande vågen åstadkommer kolsyreelimination. Amplituden ger tidalvolym som är betydligt mindre än luftvägarnas »dead space«. Oscillering genereras i olika respirator typer med specialdesignade pistoner, flödesbrytare eller elektromagnetiskt styrt membran.

Innan HFOV-tekniken introducerades i kliniskt bruk gjordes experiment på djur med surfaktantbrist som vid »respiratory distress syndrom«. I dessa studier fann man att HFOV orsakade mindre vävnadsskada än konventionell ventilation och bättre spridning av surfaktant när behandlingen påbörjades tidigt [2–5].

Den första kliniska multicenterstudien av nyfödda barn (HiFi trial) genomfördes redan innan surfaktantbehandling blev tillgänglig [6]. Studien visade inte bättre resultat, jämfört med konventionell respiratorbehandling, avseende överlevnad eller risk för utveckling av kronisk lungsjukdom (BPD), däremot ökad incidens av hjärnblödning (IVH) hos HFOV-behandlade barn. Studien kritiserades kraftfullt för att vissa deltagande centra inte behärskade tekniken med HFOV, vilket försämrade behandlingsresultatet [7]. Flera senare studier har visat bättre resultat vid HFOV på utvalda grupper utan ökad IVH-incidens [8].

Senare rapporter om användning av HFOV som förstahandsalternativ för ventilation av barn med RDS är positiva och visar mindre morbiditet och sänkta sjukvårdskostnader [9, 10]. Det är nu allmänt accepterat att HFOV är ett säkert sätt att ventilera barn födda före 34:e gestationsveckan, dock utan att man kunnat visa tydliga förbättringar i överlevnad eller sänkt förekomst av kronisk lungsjukdom, särskilt när HFOV används som »rescue«-terapi [11].

SAMMANFATTAT

HFOV (högfrekvent oscillatorventilation) för fullgångna barn med svår andningssvikt är granskad med Cochrane-metodik.

Utbredd användning av HFOV har försvårat genomförandet av randomiserade studier.

Cochrane-analysen ger en ofullständig bild av HFOV som räddningsinstrument för barn med gestationsålder >34 veckor.

Rutinmässig användning av HFOV kan inte rekommenderas för barn >34 gestationsveckor.

HFOV är en ny avancerad teknologi som kräver stora personalresurser och mycket träning.

Evidensbaserad medicin

Användning av högfrekvent oscillatorventilation av barn födda efter 34 gestationsveckor med akut hypoxisk andningssvikt sker ofta i situationer när konventionell respiratorbehandling inte ger tillfredsställande resultat, som s k »rescue«-behandling, och när patienterna har mycket hög förväntad dödlighet. I dessa situationer är ofta extrakorporeal membranoxygenering (ECMO) det andra behandlingsalternativet. ECMO ses av många som ett betydligt mer invasivt alternativ än HFOV. Målsättningen med denna översikt är att kommentera en systematisk litteraturanalys från Cochrane Collaboration avseende HFOV-behandling som »rescue«-behandling jämfört med konventionell respiratorbehandling hos fullgångna barn med svår andningssvikt.

Metaanalysen

Målsättningen med metaanalysen var att utvärdera om högfrekvent oscillatorventilation jämfört med konventionell ven-

tilation minskade mortalitet och morbiditet hos svårt lungsjuka barn över 34 gestationsveckor, utan ökning av komplikationer. Vidare hade man fördefinierat analyser av subgrupper enligt:

1. Användning av surfaktant.
2. Användning av strategier för att rekrytera hög lungvolym.
3. Användning av olika respiratorer för att leverera HFOV.
4. Barn med olika lungsjukdomar.
5. Användning av inhalerat kväveoxid tillsammans med HFOV eller konventionell respirator.

För översikten söktes slutresultat som inkluderade:

1. Dödlighet vid 28–30 dagar, före utskrivning och under första levnadsåret.
2. Användning av ECMO.
3. Respiratordagar.
4. Lung-luftläckage.
5. Syrgasbehov vid 28–30 dagars ålder och vid utskrivning.

Metaanalysens material och metoder

Cochrane Neonatal Review-gruppen har standardiserade sökmetoder för att hitta randomiserade eller kvasirandomiserade studier. Medlemmar i Cochrane Collaboration gör sökningar via Oxford Database of Perinatal Trials, Medline, Embase, tidigare översikter och korsreferenser, abstracts, konferenshandlingar, experttips och tidskriftsgenomgångar. Deltagarna arbetar på var sitt håll med att extrahera, bedöma och koda data från varje studie. Överenskommelse nås via diskussion inom gruppen. Slutdata presenteras som relativ risk (RR) och riskskillnad (RD) med 95 procent konfidensintervaller.

Endast en studie som uppfyllde sökkriterierna hittades [12]. Studien rekryterade patienter födda 1990–1992. Barn med gestationsålder >34 veckor och födelsevikt >2 000 g kunde inkluderas under de första 14 levnadsdgnen. Barnen skulle ha pågående behandling med konventionell respirator och $FiO_2 > 0,5$, medelluftvägstryck >30 cm H_2O och frekvens >40 andetag/minut. Om patienterna hade $pO_2 < 4,7$ kPa (35 mm Hg) på tre blodgasanalyser i följd, var i chocktillstånd eller behövde hjärt-lungräddning var de kvalificerade för omedelbar ECMO-behandling och uteslöts från deltagande i studien.

81 patienter randomiserades till HFOV eller konventionell respirator. Patienter som inte svarade på randomiserad behandling korsades över till motsatt behandlingsalternativ. Före randomiseringen stratifierades patienterna enligt huvuddiagnos i grupper vilka inkluderade lunginflammation, hyalinmembransjukdom, mekoniumaspirationssyndrom, luftläckage eller annat. Randomiseringen var blindad, men själva interventionen kunde självklart inte blindas. Två barn exkluderades på grund av missbildningar. Analys av ultraljudsundersökningar av skalle och diagnos av kronisk lungsjukdom var blindad mot gruppstillhörighet.

Metaanalysens resultat

Den inkluderade studien visade ingen signifikant minskning av mortalitet vid 28 dagars ålder (RR 0,51 [0,05; 5,43]) eller behandlingssvikt på den ventilationsmetoden patienten först randomiserades till (RR 0,73 [0,47; 1,13]). Det var ingen signifikant skillnad i antalet patienter som behövde ECMO-behandling (RR 2,05 [0,85; 4,92]) eller antal respiratordagar, tid med extra syrgastillförsel och sjukhusdagar. Studien redovisade ingen långtidsuppföljning av tillväxt eller utveckling hos patienterna.

Metaanalysens diskussion och konklusioner

Endast en studie uppfyllde kriterier för att inkluderas i metaanalysen. I diskussionen påpekar man att studien hade in-

byggda kriterier för korsning mellan de två ventilationsmetoderna och att detta ledde till en underskattning av eventuella behandlingseffekter. Dessutom beräknade man ett behov av att randomisera 250 patienter för att påvisa signifikanta skillnader mellan de två ventilationsmetoderna. Studien avbröts i förtid på grund av rekryteringsproblem beroende på ökande användning av HFOV på inremitterande sjukhus. Studien kunde därför inte svara på den ställda frågan.

Författarnas konklusion är därför att det inte finns randomiserade kontrollerade data som stödjer rutinmässig användning av högfrekvent oscillatorventilation av barn >34 gestationsveckor med svår andningssvikt. Med hänsyn till den ökande användningen av HFOV påpekar de att det är brådskande att ta fram mer data som genereras via randomiserade kontrollerade studier. Detta för att bekräfta HFOV-behandlingens roll. Vidare att i eventuella framtida studier blir uppdelning enligt grunddiagnos och noggrann uppföljning av patienterna viktig.

II Kommentar

Högfrekvent oscillatorventilation av nyfödda barn har nu passerat sitt första decennium i kliniskt bruk. De djurexperiment som gjordes innan metoden började användas på människa var extremt lovande. Alla använde sig av modeller som efterliknar hyalinmembransjukdom med homogent utspridda förändringar som gav reducerad volym i båda lungorna. Tekniken visade sig sedan ge upphov till mindre lungskador i djurexperimenten och kunde främja lungans egen produktion av surfaktant [13].

Den första stora multicenterstudien, som publicerades 1993, gav däremot upphov till bekymmer för läkare som ville använda tekniken på nyfödda barn. Det visade sig nämligen att resultaten inte skiljde sig nämnvärt från resultat som uppnås med konventionell respirator, och man fick en ökad förekomst av hjärnblödningar hos barnen som behandlades med högfrekvent oscillatorventilation. Studien kritiserades kraftfullt av läkare som utvecklade tekniken för att man inte följde de högvolymstrategier som ligger bakom den goda effekten på syresättningen. Nämligen att »öppna lungan och hålla den öppen«, vilket är den grundläggande principen bakom syresättning vid HFOV [14]. I studien som är gjord på över 50 olika kliniker fanns många som inte behärskade tekniken. Följderna blev att behandlingsresultaten försämrades.

Man har från början talat om att HFOV-tekniken är bäst lämpad för lungsjukdomar med generellt sänkt lungvolym och mindre passande för lungsjukdomar som har icke homogen patologi, exempelvis tidig mekoniumaspiration, pneumoni eller diafragmabräck. Eftersom hyalinmembransjukdom är mycket sällsynt hos barn efter 34 gestationsveckor är det ofta icke homogen lungpatologi som ligger bakom HFOV-kandidaternas svåra andningssvikt. Detta syns i hittills publicerade studier på effekterna av HFOV och gör resultaten svårtolkade. Dessutom är det ofta frågan om »cross over«-design på studierna som gör det svårt att tolka behandlingsresultaten, exempelvis den aktuella metaanalysen. Under senaste decenniet har andra behandlingsalternativ börjat användas inom neonatalmedicinen, exempelvis surfaktantbehandling, inhalerad kväveoxid och vasoaktiva mediciner. En randomiserad studie använde kväveoxidterapi tillsammans med HFOV. Studien visar att högfrekvent oscillatorventilation, genom att rekrytera lungvolym, förbättrar effekten av den inhalerade kväveoxiden [15].

Det är designen på studierna, det heterogena patientmaterialet och blandning med annan terapi som komplicerar utvärderingen av effekten av HFOV hos barn >34 veckor. HFOV-behandling används i dag på de flesta regionscentra för nyfödda med svår andningssvikt och har tyvärr kommit i kliniskt



Högfrekvent oscillatorventilation för fullgångna barn med svår andningssvikt är en ny avancerad teknologi som kräver stora personalresurser och mycket träning. Trots den utbredda användningen behövs nya multicenterbaserade randomiserade studier, i vilka det är nödvändigt att stratifiera patienterna enligt underliggande lungpatologi och kompletterande behandlingsstrategier. Tillgång till ECMO-terapi bör finnas där studierna genomförs. Detta är nödvändigt för att fastställa HFOV-behandlingens definitiva plats i behandlingskedjan.

bruk innan den definitiva randomiserade forskningen gjordes. Liknande forskningssituation gällde när konventionell respiratorbehandling av för tidigt födda barn började användas i kliniken [16]. ECMO-behandling är ofta sista instans för barn med svår andningssvikt.

Ny sammanställning av 12 175 fall från den världsomfattande ECMO Life Support Organization (ELSO) visar att under åren 1988–1998 har man ökat från ingen användning av surfaktant till 36 procent, HFOV till 46 procent och NO till 24 procent innan patienterna hamnar i ECMO [17]. ECMO-terapi på barn med mekoniumaspiration har minskat när man använder sig av HFOV- och NO-behandling i kombination [15, 18]. Det blir därför svårt att ta ett steg tillbaka i den kliniska forskningen eftersom det i praktiken blir omöjligt att rekrytera patienter.

Svenska erfarenheter av HFOV-behandling

Neonatalavdelningen på Astrid Lindgrens barnsjukhus, Karolinska sjukhuset, utgör regioncentrum för den neonatala intensivvården i Stockholms läns landsting (22 000 förlösningsår). Neonatalavdelningen på Karolinska sjukhuset tog HFOV som »rescue«-åtgärd i kliniskt bruk 1993 som första klinik i Sverige. Under perioden augusti 1993 till december 1998 respiratorbehandlades 471 barn. Av dessa barn behövde 87 (18 procent) behandling med HFOV. Alla barn hade trots maximal konventionell respiratorbehandling inadekvat syresättning och/eller dålig ventilation enligt kriterier som i den aktuella metaanalysen. Överlevnaden hos HFOV-behandlade barn var 60/87 (69 procent). HFOV-behandling påbörjades i genomsnitt vid 45 timmars ålder, och patienterna behandlades i genomsnitt i 57 timmar. Av 60 överlevande barn kunde 55 (91 procent) extubereras från HFOV direkt till nasal CPAP-behandling. Intraventrikulär blödning (IVH) hittades hos nio patienter, med fyra blödningar av grad III–IV. Blödningen ansågs vara direkt bidragande dödsorsak hos två av dessa barn. Båda var extremt prematura.

Det som skiljer vår erfarenhet av högfrekvent oscillatorventilation från andra centras är att nästan alla våra HFOV-patienter kunde extubereras direkt till nasalt CPAP. Detta förkortar patientens totala respiratortid och har därför klara fördelar jämfört med en mellanlandning i konventionell respirator. Komplikationer i samband med behandlingen hos dessa

patienter har varit överexpansion av lungor som är under tillfrisknande samt kardiovaskulära problem. Från 1993 har våra resultat förbättrats beroende på dels ökad användning av högfrekvent oscillatorventilation, dels ökad kunskap om utrustningen och behandlingsprinciperna.

II Konklusion

Den aktuella metaanalysen ger en ofullständig bild av HFOV-behandling som räddningsinstrument för barn med gestationsålder >34 veckor. Studien som den baseras på var för liten för att svara på den ställda frågan. Studier på dessa patienter försvåras på grund av att de ofta har olika lungsjukdomar som orsak till sin svåra andningssvikt, och att de därför blir svåra att jämföra. Ny adjungerande terapi har tillkommit, exempelvis surfaktant och NO(kväveoxid)-behandling som ofta används tillsammans med HFOV-behandlingen och förbättrar resultaten. Högfrekvent oscillatorventilation har blivit så spridd att bara detta faktum i sig försvårar ytterligare studier.

Metaanalysens slutsats när det gäller forskning kan man dock hålla med om. Den innebär att trots den utbredda användningen av HFOV behövs nya multicenterbaserade randomiserade studier. I dessa studier är det nödvändigt att stratifiera patienterna enligt underliggande lungpatologi och kompletterande behandlingsstrategier. Tillgång till ECMO-terapi bör finnas där dessa studier genomförs. Detta är nödvändigt för att fastställa HFOV-behandlingens definitiva plats i behandlingskedjan.

Referenser

1. Nilsson R. The artificially ventilated preterm rabbit neonate as experimental model of hyaline membrane disease. *Acta Anaesth Scand* 1982;26:89-103.
2. McCulloch PR, Forkert P, Froese AB. Lung volume maintenance prevents lung injury during high frequency oscillatory ventilation in surfactant deficient rabbits. *Am Rev Respir Dis* 1988;137:1185-92.
3. Froese AB, McCulloch PR, Sugiura M, Vaclavik S, Possmayer F, Moller F. Optimizing alveolar expansion prolongs the effectiveness of exogenous surfactant therapy in the adult rabbit. *Am Rev Respir Dis* 1993;148:569-77.
4. Meredith KS, deLemos RA, Coalson JJ, King RJ, Gerstmann DR, Kumar R, et al. Role of lung injury in the pathogenesis of hyaline membrane disease in premature baboons. *J Appl Physiol* 1989;66(5):2150-8.
5. Jackson JC, Truog WE, Standaert TA, Murphy JH, Juul SE, Chi EY, et al. Reduction in lung injury after combined surfactant and high-frequency ventilation. *Am J Respir Crit Care Med* 1994;150:534-9.
6. HiFi Study Group HS. High-frequency oscillatory ventilation compared with conventional mechanical ventilation in the treatment of respiratory failure in preterm infants. *N Engl J Med* 1989;320:88-93.
7. Bryan A, Froese AB. Reflections on the HiFi trial. *Pediatrics* 1991;87:565-7.
8. Clark R, Dykes F, Bachman T, Ashurst J. Intraventricular hemorrhage and high frequency ventilation: a meta-analysis of prospective clinical trials. *Pediatrics* 1996;98:1058-61.
9. Gerstmann DR, Minton SD, Stoddard RA, Meredith KS, Monaco F, Bertrand JM, et al. The Provo multicenter early high-frequency oscillatory ventilation trial: improved pulmonary and clinical outcome in respiratory distress syndrome. *Pediatrics* 1996;98(6):1044-57.
10. Clark R. Prospective randomized comparison of high-frequency oscillatory and conventional ventilation in respiratory distress syndrome. *Pediatrics* 1992;89:5-12.
11. Bhuta T, Henderson-Smart DJ. Rescue high frequency oscillatory ventilation versus conventional ventilation for pulmonary dysfunction in preterm infants (Cochrane Review). In: *The Cochrane Library*, 4, 2001. Oxford: Update Software.
12. Clark R, Yoder B, Sell M. Prospective, randomized comparison of high-frequency oscillation and conventional ventilation in candidates for extracorporeal membrane oxygenation. *J Pediatr* 1994;124:447-54.

13. de Lemos RA, Coalson JJ, Gerstmann DR, Null DM Jr, Ackerman NB, Escobedo MB, et al. Ventilatory management of infant baboons with hyaline membrane disease: the use of high frequency ventilation. *Pediatr Res* 1987;21:594-602.
14. Froese A, Bryant A. State of Art: High Frequency Ventilation. *Am Rev Respir Dis* 1987;135:1363-74.
15. Kinsella JP, Truog WE, Walsh WF, Goldberg RN, Bancalari E, Mayock DE, et al. Randomized multicenter trial of inhaled nitric oxide and high frequency oscillatory ventilation in severe, persistent pulmonary hypertension of the newborn. *J Pediatr* 1997;131:55-62.
16. Bancalari E, Sinclair J. Mechanical ventilation, in effective care of the newborn infant. Sinclair J, Bracken M, editors. Oxford: Oxford University Press; 1992. p. 200-18.
17. Roy B, Rycus P, Conrad S, Clark R. The changing demographics of neonatal ECMO patients reported to the ELSO registry. *Pediatrics* 2000;106:1334-8.
18. Wilson J, Bower L, Thompson J, Fauza D, Fackler J. ECMO in evolution: the impact of changing patient demographics and alternative therapies on ECMO. *J Pediatr Surg* 1996;31:1116-22.

SUMMARY

Commentary on a Cochrane review
High frequency oscillatory ventilation (HFOV) or conventional mechanical ventilation in severe respiratory failure for newborns more than 34 weeks gestational age

Baldvin Jónsson

Läkartidningen 2001;98:5626-9

The Cochrane review of the use of high frequency oscillatory ventilation in newborns more than 34 weeks gestation only included one randomized study and therefore gives an incomplete picture of the use of HFOV as rescue therapy for this patient group. The study was stopped prematurely due to difficulties in patient recruitment. There is a real need for more randomized studies but increasing use of HFOV has made these studies very difficult to perform. From the existing data routine use of HFOV in this patient group cannot be recommended.

Experience of using HFOV in one Swedish tertiary care center during the years 1993-1998 is described. Eighty seven newborns were treated with HFOV rescue with 69 percent patient survival while 91 percent of the surviving infants were extubated directly to nasal continuous positive airway pressure. Intraventricular hemorrhage was seen in 9 patients.

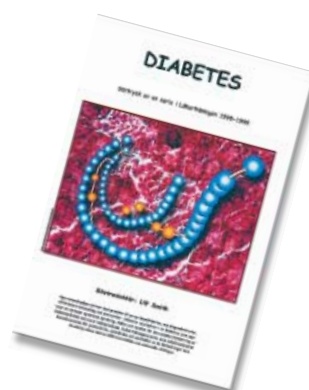
Correspondence: Baldvin Jónsson, Neonatalprogrammet, Astrid Lindgrens barnsjukhus, Karolinska sjukhuset, SE-171 76 Stockholm, Sweden. (Baldvin.Jnsson@kbh.ki.se)

Särtryck

Läkartidningen

Nya vetenskapliga rön har lagt grunden till en ny klassifikation, nya diagnoskriterier, effektivare behandling och prevention – inklusive vaccination – av diabetes, som uppvisar en närmast epidemisk spridning. Målen och medlen för den snabba förbättringen av diabetesvården som nu är möjlig belyses i Läkartidningens serie, som också analyserar konsekvenserna för patienterna, sjukvården och samhället av de förändringar som Socialstyrelsen hösten 1999 fastställde som nationella riktlinjer.

Priset är 85 kronor.



Diabetes

Beställer härmed.....ex
av "Diabetes"

.....
namn

.....
adress

.....
postnummer

.....
postadress

**Insändes till Läkartidningen
Box 5603
114 86 Stockholm**

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker

Noggrann anamnes kan ge besked om typ av primär huvudvärk

Redovisning av och kommentarer till omfattande översiktsartikel

■ Tillkomsten av ett nytt internationellt klassifikationssystem, tillsammans med imponerande forskningsinsatser under de senaste årtiondena, har dramatiskt ökat möjligheterna till att korrekt kunna diagnostisera och vid behov behandla olika smärttillstånd utlösta inom huvudregionen. Det nya förslaget till klassifikation och diagnostiska kriterier av huvudvärks-åkommor publicerades 1988 i form av ett supplement i *Cephalalgia* [1]. Från de tidigare alltför diffust uppfattade huvudvärkssyndromen kan man idag urskilja ett stort antal specifika sjukdomstillstånd som ofta kräver särskild behandling och/eller omhändertagande. Den nya internationella huvudvärkssklassifikationen består av 13 olika grupper, varav de 4 första skall täcka in de primära huvudvärksformerna och de övriga de sekundära (Faktaruta 1). Diagnostisering av primära huvudvärksformer baserar sig helt och hållet på de anamnestiska uppgifterna.

Huvudvärkssjukdomar

Två huvudvärkssjukdomar är betydligt vanligare än alla övriga, nämligen migrän och huvudvärk av spänningstyp. Dessa två huvudvärksformer står för över 90 procent för alla som söker sjukvård med just huvudvärk som dominerande symtom. Migrän delas primärt in i två undergrupper, migrän utan respektive med aura. Migrän utan aura (tidigare använda beteckningar är vanlig migrän eller hemicrania simplex) beskrivs enligt följande: idiopatiskt, attackvis återkommande huvudvärkssjukdom som manifesteras i attacker med 4–72 timmars duration. Huvudvärken kännetecknas av ensidig lokalisering, pulserande karaktär, måttlig till svår intensitet, försämring av normal fysisk aktivitet samt tillhörande symtom som illamående, ljus- och ljudkänslighet (Faktaruta 2).

Migrän med aura (tidigare använda beteckningar är klassisk, oftalmisk, hemiparetisk, hemiplegisk eller afatisk migrän, migrän accompagnée eller komplicerad migrän) beskrivs enligt följande: idiopatiskt, attackvis återkommande huvudvärkssjukdom som manifesteras i attacker med neurologiska symtom, vilka otvetydigt kan lokaliseras till hjärnbarken eller hjärnstammen, vanligtvis utvecklas under loppet av 5–20 minuter och sällan pågår längre än 60 minuter. Huvudvärk, illamående och/eller ljuskänslighet följer vanligtvis direkt efter de neurologiska symtomen eller inom ett fritt intervall på mindre än en timme. Huvudvärken varar

SAMMANFATTAT

Primära huvudvärksformer är kliniska diagnoser som närmast uteslutande baseras på anamnes. Somatiska och neurologiska undersökningar kan dock ge viss försäkran om att huvudvärken inte är orsakad av sekundära huvudvärksformer.

Existerande kriterier för klassifikation och diagnostisering av huvudvärk, enligt internationella rekommendationer, upplyser inte om vilka kliniska symtom och egenskaper som är mest betydelsefulla för att fastställa eller utesluta primära huvudvärksformer.

Baserat på tillgängliga engelskspråkliga artiklar och tillhörande bibliografi januari 1966–maj 1999 har Smetana beräknat sensitivitet och specificitet för kliniska symtom och egenskaper, som ett mått på det diagnostiska värdet av anamnestiska uppgifter vid primära huvudvärksformer.

De kliniska egenskaper som är mest prediktiva för migrän i jämförelse med huvudvärk av spänningstyp är illamående, fotofobi, fonofobi och intensifierad huvudvärk/smärta vid fysisk ansträngning samt aura.

Dietära triggerfaktorer är mer prediktiva för migrän. Andra utlösande faktorer är lika vanliga vid migrän och huvudvärk av spänningstyp.

Hortons huvudvärk (klusterhuvudvärk) skiljer sig från migrän och huvudvärk av spänningstyp med uppgifter om strikt halvsidighet, intensiv periorbital smärta och attackduration på mindre än en timma hos de flesta individer.

Evidensbaserad medicin

II Fakta 1

Primära och sekundära huvudvärksformer enligt International Headache Society

Primära huvudvärksformer

1. Migrän
2. Huvudvärk av spänningstyp
3. Hortons huvudvärk och kronisk paroxysmal hemikrani
4. Övrig huvudvärk – ej associerad med vävnadsskada

Sekundära huvudvärksformer

5. Huvudvärk associerad med trauma mot huvudet
6. Huvudvärk associerad med vaskulär sjukdom
7. Huvudvärk associerad med icke vaskulär intrakraniell sjukdom
8. Huvudvärk associerad med substanser eller utsättande av substanser
9. Huvudvärk associerad med icke-cefal infektion
10. Huvudvärk associerad med metabolisk sjukdom
11. Huvudvärk eller facial smärta associerad med sjukdom i kraniet, nacken, ögon, öron, näsa, bihålor, tänder, mun eller andra ansikts- eller kranieala strukturer
12. Kranieala neuralgier, nervtrådsmärta och deafferentations smärta
13. Icke klassificerbar huvudvärk

vanligtvis mellan 4 och 72 timmar, men kan saknas helt (Faktaruta 2).

Huvudvärk av spänningstyp delas primärt in i två undergrupper, episodisk form respektive kronisk form (minst 15 dagar varje månad under minst 6 månader). Episodisk huvudvärk av spänningstyp (tidigare använda beteckningar är spänningshuvudvärk, psykomogen huvudvärk, stresshuvudvärk, vanlig, essentiell, idiopatisk eller psykogen huvudvärk) beskrivs enligt följande: återkommande episoder av huvudvärk som varar från minuter till flera dygn. Huvudvärken är i typiska fall tryckande eller åtstramande till karaktären, med lätt till måttlig svårighetsgrad, dubbelsidig lokalisation och utan försämring vid normal fysisk aktivitet. Yrsel saknas, men ökad ljus- eller ljudkänslighet kan förekomma (Faktaruta 2).

Hortons huvudvärk delas primärt in i två undergrupper, episodisk respektive kronisk form.

Hortons huvudvärk (tidigare använda beteckningar är Bings erytroprosopalmi, ciliarisneuralgi, migränös neuralgi (Harris), erytromelalmi i huvudet, histaminhuvudvärk, petrosusneuralgi (Gardner), sfenopalatinusneuralgi, Vidians eller Sluders neuralgi eller hemikrania periodica neuralgiformis) beskrivs enligt följande: Attackerna av svår, strikt ensidig smärta belägen i eller ovanför ögat och/eller tinningen, som varar mellan 15 minuter och 3 timmar, och som kan uppträda från 1 gång varannan dag till 8 gånger under samma dygn. Under attackerna förekommer ett eller flera av följande symtom: blodkärlfyllnad i ögats bindhinna, tårflöde, nästäppa, rinnsnuva, pann- och ansiktssvettning, liten pupill, nedsjunket ögonlock samt ögonlockssvullnad. Attackerna uppträder i serier som varar från veckor till månader (så kallade Hortonperioder), vilka är åtskilda av remissionsperioder (smärtfria intervall) som vanligtvis varar månader till år. Hortonperiodernas längd är vanligtvis mellan två veckor och tre månader. Hos cirka 10 procent av patienterna är förloppet kroniskt, utan remissionsperioder (Faktaruta 2).

Trots den nya huvudvärksklassifikationen kan migrän och

huvudvärk av spänningstyp ibland vara svåra att skilja åt, såväl för patienten som den diagnostiserande läkaren. En patient kan dessutom ha flera olika former av huvudvärk samtidigt. Inte sällan förekommer migrän och huvudvärk av spänningstyp hos en och samma patient [2]. Än idag saknas objektiva test för de flesta huvudvärksjukdomar. Det finns inte några specifika laboratorieundersökningar, röntgenologiska, neurofysiologiska eller några andra metoder som kan användas för att påvisa sjukdomsorsaken. Läkaren är därför i mycket hög grad beroende av patientens egen sjukdomsbeskrivning vid huvudvärksdiagnostik. En grundlig genomgång av en patient som söker på grund av huvudvärk tar tid, inte sällan mer än 30 minuter. Huvuddelen av denna tid bör ägnas åt en grundlig huvudvärksanamnes då resultatet av klinisk somatisk/neurologisk undersökning oftast är av underordnad betydelse vid huvudvärksdiagnostik. I allmänhet kan läkaren efter avslutad intervju och undersökning ställa en preliminär diagnos enligt den nya huvudvärksklassifikationen eller bedöma om ytterligare kompletterande undersökningar behövs för att kunna ställa rätt diagnos. Det som ofta driver en patient med huvudvärk till sjukvården är fruktan för en bakomliggande hjärntumör. Även om hjärntumör är en relativt ovanlig orsak till huvudvärk så motiverar inte minst denna rädsla hos huvudvärkspatienten att anamnesupptagandet och statusgenomgången får ta sin väl tilltagna tid. Det är psykologiskt oerhört viktigt att en huvudvärkspatient kan gå hem med klara lugnande besked. En noggrant tagen huvudvärksanamnes minskar markant behovet av ytterligare utredning för att utsluta sekundära huvudvärksformer.

Översiktsartikelns bakgrund och metod

Utgångsläget för Smetanas översiktsartikel i Archives of Internal Medicine är att den nya huvudvärksklassifikationen förefaller vara ett bra verktyg för huvudvärksspecialister, men kanske ett alltför komplicerat instrument för att användas i primärvårdsläkarens dagliga kliniska arbete. Vidare att det av huvudvärksklassifikationen inte framgår vilka av de föreslagna kliniska symtomen/egenskaperna som är mest värdefulla för differentiering mellan och diagnostisering av olika primära huvudvärksformer.

Målsättningen med artikeln var att bestämma det diagnostiska värdet av anamnestiska uppgifter vid primära huvudvärksformer.

Författaren gjorde en sökning på MEDLINE av engelskspråkiga artiklar som publicerats 1 januari 1966 till 31 maj 1999, varav 85 stycken finns medtagna i författarens referenslista. De sökord som användes inkluderade migrän (migraine), huvudvärk av spänningstyp (tension-type headache), Hortons huvudvärk (cluster headache), kliniska egenskaper (clinical features), symtom (symptoms), anamnes (history), aura (aura), utlösande faktorer (precipitants), triggerfaktorer (trigger), ärftlighet (family history) och åksjuka (motion sickness). Ytterligare artiklar identifierades vid genomgång av tidigare översiktsartiklar, läroböcker och monografier. Artiklarna klassades efter vilka huvudvärkskriterier som använts vid diagnostisering (den nya huvudvärksklassifikationen eller äldre traditionella metoder). Av naturliga skäl kunde författaren vid genomgång av artiklar omfattande en tidsperiod på drygt 33 år inte använda sig av någon enhetlig och specifik diagnostisk standard.

Smetana bestämde att enskilda symtom och anamnestiska uppgifter i de granskade artiklarna skulle utgöra det diagnostiska testet för de primära huvudvärksformerna migrän, huvudvärk av spänningstyp och Hortons huvudvärk [3]. Sensitivitet definierades som andelen patienter som hade en given primär huvudvärksdiagnos med symtomet/egenskapen ifråga. Specificitet definierades som andelen patienter utan den

II Fakta 2

Diagnoskriterier enligt International Headache Society

Migrän utan aura (IHS 1.1)

- A. Åtminstone fem attacker som uppfyller kriterierna B–D krävs.
- B. Huvudvärksattacker som pågår under 4–72 timmar (obehandlade eller icke framgångsrikt behandlade).
- C. Huvudvärk som har åtminstone två av följande kännetecken: 1. Ensidig lokalisering. 2. Pulserande karaktär. 3. Måttlig till svår intensitet (hämmar eller förhindrar dagliga aktiviteter). 4. Försämring av gång i trappor eller motsvarande normal fysisk aktivitet.
- D. Under huvudvärken åtminstone ett av följande: 1. Illamående och/eller kräkningar. 2. Ljus känslighet och ljudkänslighet.
- E. Åtminstone ett av följande: 1. Sjukhistoria, allmän somatisk och neurologisk undersökning tyder inte på något av de sjukdomstillstånd som listas i grupperna 5-11 (Faktaruta 1). 2. Sjukhistoria och/eller allmän somatisk och/eller neurologisk undersökning tyder på ett sådant sjukdomstillstånd, men detta är uteslutet genom lämpliga undersökningar. 3. Ett sådant sjukdomstillstånd finns, men migränattackerna uppträder inte för första gången i nära tidsmässigt samband med detta tillstånd.

Migrän med aura (IHS 1.2)

- A. Åtminstone två attacker som uppfyller kriteriet B krävs.
- B. Åtminstone tre av följande fyra kännetecken: 1. Ett eller flera helt reversibla aurasymtom som tyder på fokala hjärnbarks- och/eller hjärnstamsdysfunktion. 2. Åtminstone ett aurasymtom som utvecklas gradvis under mer än fyra minuter eller två eller flera symtom som uppträder efter varandra. 3. Inget enskilt aurasymtom varar mer än 60 minuter. Om mer än ett aurasymtom förekommer kan durationen vara proportionellt ökad. 4. Huvudvärk följer efter auran med ett fritt intervall på mindre än 60 minuter (den kan också börja före eller samtidigt med aurasymtomen).
- C. Åtminstone ett av följande: 1. Sjukhistoria, allmän somatisk och neurologisk undersökning tyder inte på något av de sjukdomstillstånd som listas i grupperna 5-11 (Faktaruta 1). 2. Sjukhistoria och/eller allmän somatisk och/eller neurologisk undersökning tyder på ett sådant sjukdomstillstånd, men detta är uteslutet genom lämpliga undersökningar. 3. Ett sådant sjukdomstillstånd finns, men migränattackerna uppträder inte för första gången i nära tidsmässigt samband med detta tillstånd.

Episodisk huvudvärk av spänningstyp (IHS 2.1)

- A. Minst tio tidigare episoder av huvudvärk med följande karaktär, B–D:
- B. Duration 30 minuter–7 dygn (högst 180 dagar/år).
- C. Minst två av följande smärtekaraktäristika: 1. Molande, tryckande (inte pulserande). 2. Lätt till måttlig (kan hämma men inte hindra aktiviteter). 3. Bilateral lokalisering (dubbelsidig). 4. Förvärras inte av trappgång eller måttlig fysisk aktivitet.
- D. Bägge av följande: 1. Inget illamående eller kräkningar (matleda får förekomma). 2. Avsaknad av fotofobi och fonofobi (ettdera får förekomma).
- E. Inte orsakad av somatisk, kranio cerebral sjukdom eller skada.

Hortons huvudvärk med obestämd periodicitet (IHS 3.1)

- A. Åtminstone fem attacker som uppfyller kriterierna B–D.
- B. Svår ensidig smärta i ögat, ovanför ögat och/eller vid tinningen, vilken obehandlad varar mellan 15 och 180 minuter.
- C. Huvudvärken är förenad med åtminstone ett av följande kännetecken, som måste finnas på smärtsidan: 1. Konjunktival kärlinjektion. 2. Tårflöde. 3. Nästäppa. 4. Snuva. 5. Svette i panna och ansikte. 6. Sammandragen pupill. 7. Nedsjunktet ögonlock. 8. Svullnad i ögonlock.
- D. Attackfrekvensen varierar från en varannan dag till åtta per dygn.
- E. Åtminstone ett av följande: 1. Sjukhistoria, allmän somatisk och neurologisk undersökning tyder inte på något av de sjukdomstillstånd som listas i grupperna 5-11 (Faktaruta 1). 2. Sjukhistoria och/eller allmän somatisk och/eller neurologisk undersökning tyder på ett sådant sjukdomstillstånd, men detta kan uteslutas genom lämpliga undersökningar. 3. Ett sådant sjukdomstillstånd finns, men Hortons huvudvärk uppträder inte för första gången i nära tidsmässigt samband med detta tillstånd.

Episodisk Hortons huvudvärk (IHS 3.1.2)

Uppträder i perioder som varar mellan sju dygn och ett år, med mellanliggande remissionsperioder som varar 14 dygn eller längre.

- A. Uppfyller kriterierna för 3.1.
- B. Minst två perioder av huvudvärk (Hortonperioder) som (hos obehandlade patienter) varar från sju dygn till ett år och är åtskilda av remissionsperioder (smärtfria intervall) som varar 14 dygn eller längre.

primära huvudvärksdiagnosen som inte hade symtomet/egenskapen ifråga. För migrän och huvudvärk av spänningstyp förelåg tillräckligt många observationer för att beräkna den positiva respektive negativa sannolikhetskvoten för kliniska egenskaper/symtom och för utlösande faktorer [4, 5]. En positiv kvot, sensitivitet/(1 – specificitet), indikerar en ökad sannolikhet för en viss primär huvudvärksdiagnos om en specifik klinisk egenskap eller ett specifikt symtom finns närvarande. En negativ kvot, (1 – sensitivitet)/specificitet, indikerar en minskad sannolikhet för en viss primär huvudvärks-

diagnos om specifik klinisk egenskap/specifikt kliniskt symtom inte finns närvarande. Konfidensintervall beräknades enligt metod beskriven av Simel och medarbetare [4].

Huvudresultaten

Andelen patienter med diagnosen migrän, enligt International Headache Society (IHS), som hade symtomen/egenskaperna (sensitivitet) ensidig huvudvärk, pulserande karaktär, måttlig till svår intensitet, försämring av normal fysisk aktivitet samt tillhörande symtom som illamående, ljus- och ljudkänslighet

Tabell IV. Sensitivitet och specificitet för symtom/egenskaper för migrän enligt International Headache Society jämfört med huvudvärk av spänningstyp samt beräknad positiv sannolikhetsknot [1].

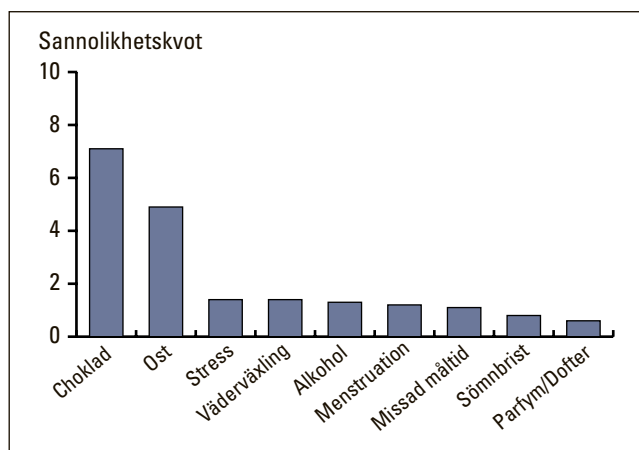
Symtom/egenskap	Sensitivitet	Specificitet	Sannolikhetsknot
Illamående	82 procent	96 procent	23,2
Ljuskänslighet	79 procent	87 procent	6,0
Ljudkänslighet	69 procent	87 procent	5,2
Huvudvärk förstärks av fysisk aktivitet	81 procent	78 procent	3,7
Dunkande/pulserande	76 procent	77 procent	3,3
Halvsidig huvudvärk	66 procent	78 procent	3,1
Duration 4–24 tim.	57 procent	67 procent	1,7
Duration 24–72 tim.	13 procent	91 procent	1,4

anges i Tabell I. I samma tabell anges även andelen patienter utan migrän men med huvudvärk av spänningstyp enligt IHS som inte hade symtomen/egenskaperna ifråga (specificitet).

Mest sensitiva för diagnosen migrän var illamående, försämring av huvudvärken vid normal fysisk aktivitet, ljuskänslighet och dunkande/pulserande kvalitet. Dessa angavs av 82, 81, 79 respektive 76 procent av dem med migrän. Sensitiviteten för ljudkänslighet och halvsidighet var lägre och angavs enbart av 69 respektive 66 procent av de med diagnosen migrän. Illamående, ljus- och ljudkänslighet var de mest specifika symtomen då 96, 87 respektive 87 procent av de med huvudvärk av spänningstyp inte hade dessa symtom. Försämring av huvudvärken vid fysisk ansträngning, den dunkande/pulserande kvaliteten liksom halvsidigheten hade något lägre specificitet, 78, 77 respektive 78 procent av dem med huvudvärk av spänningstyp hade inte dessa symtom. Dessa procentsatser på sensitivitet och specificitet användes för beräkning av den positiva sannolikhetsknoten för diagnosen migrän. Förekomst av illamående (kvot 23,2), ökad känslighet för ljus (kvot 6,0) och ljud (kvot 5,2) talar mest för migrän (Tabell I). Försämring av huvudvärken vid normal fysisk aktivitet (kvot 3,7), dunkande/pulserande kvalitet (kvot 3,3) och halvsidighet (kvot 3,1) var förenade med en något mindre sannolikhet för diagnosen migrän. Intressant nog visade sig den anamnestiska uppgiften om ärftlighet (förekomst av migrän hos nära anhörig) vara förenad med en sensitivitet på 58 procent och en specificitet på 88 procent. Det ger en beräknad positiv sannolikhetsknot på 5,0, vilket gör denna information till en värdefull differentialdiagnostisk uppgift. Betydligt starkare än anamnestiska uppgifter om kräkningsattacker i barndomen (kvot 2,4) och förekomst av åksjuka (kvot 2,2).

Den negativa sannolikhetsknoten för diagnosen migrän var mindre än eller lika med 0,4 för illamående (kvot 0,2), ökad känslighet för ljus (kvot 0,2) och ljud (kvot 0,4), försämring av huvudvärken vid normal fysisk aktivitet (kvot 0,2) och dunkande/pulserande kvalitet (kvot 0,3). Frånvaro av något av dessa symtom talar således mot diagnosen migrän.

Vad gäller utlösande faktorer föreföll de flesta (stress, menstruation, missad måltid, sömnbrist och starka dofter) vara gemensamma för migrän och huvudvärk av spänningstyp, och besatt därför inte något differentialdiagnostiskt värde. Möjligen utgör dietära triggerfaktorer och väderomslag undantag eftersom majoriteten (choklad 95 procent, ost 92 procent, alkohol 77 procent och väderomslag 74 procent) av individerna med huvudvärk av spänningstyp inte angav dessa som huvudvärksutlösande faktorer. Choklad var den mest specifika triggerfaktorn med en beräknad positiv sannolikhetsknot på 7,1 (motsvarande för IHS-diagnostiserade 4,6)



Figur 1. Positiv sannolikhetsknot för diagnosen migrän jämfört med huvudvärk av spänningstyp [3].

för migrändiagnos när individer med migrän jämfördes med individer med huvudvärk av spänningstyp (Figur 1).

Författaren konkluderar att den mest sensitiva egenskapen för migrändiagnos är illamående, förstärkning av huvudvärken vid normal fysisk aktivitet och ökad känslighet för ljus. De mest specifika symtomen för att särskilja migrän från huvudvärk av spänningstyp är illamående, ökad känslighet för ljus och för ljud. Bland aurasymtomen är de visuella fenomenen vanligast, och 84 procent av de med migrän med aura angav denna typ av aurasymtom. Durationen av auran är vanligen kortare än 30 minuter. Migränsjukdomen är tre gånger vanligare bland kvinnor, och mer än hälften uppvisar ärftlighet för migrän. Slutligen drar författaren slutsatsen att Hortons huvudvärk (klusterhuvudvärk) är strikt halvsidig, periorbital hos 80 procent av patienterna och med attacker som varar mindre än en timme hos 54 procent av patienterna.

Kommentar om arbetets betydelse

Den kliniska kunskapen om migränanfallets naturlförlopp är mycket begränsad då de flesta patienter i dagsläget använder symtomlindrande läkemedel. Mycket sällan låter individen anfallet fortgå obehandlat. Det kan därför vara svårt att uttala sig om vad som är äkta migränsymtom under anfallet, och att skilja biverkningar av läkemedelsbehandling från kvarstående symtom av anfallet. Det kliniska intrycket från individer som besväras av migränsjukdomen är att de under prodromalfasen, som kan föregå själva migränanfallet med mer än ett dygn, kan känna mer eller mindre uttalad trötthet, koncentrationssvårigheter, tankeblockad, ökad irritabilitet, begär efter vissa födoämnen, matleda, ökad törst med mera. I frånvaro av aurasymtom börjar därefter migränanfallet med en halvsidig, dunkande/pulserande huvudvärk av hög intensitet, vilken dessutom förvärras av fysisk ansträngning, skakning av huvudet eller Valsalvamanöver. Parallellt med huvudvärken kommer ökad känslighet för ljus och ljud (eventuellt även lukt) samt illamående, vilket kan kulminera i kräkning. Om denna kliniska bild styrks av typiska aurasymtom, periodicitet med normalt status mellan anfallen och en ärftlig belastning föreligger knappast någon tvekan om migrändiagnosen.

Huvudvärk av spänningstyp skiljer sig i flera avseenden från migrän, bland annat genom att det upplevs mer som tryckande eller åtstramande, som ett »pannband runt huvudet« eller en »blymossa«. Värken är normalt inte pulserande och förvärras inte heller av fysisk aktivitet. Huvudvärk av spänningstyp kommer under pågående stress, medan migrän oftast kommer efter stressituationen när man försöker koppla av. Den blir ofta, till skillnad från migrän, bättre under mo-

tionsaktiviteter. Det är emellertid inte ovanligt att patienten, trots lång personlig erfarenhet, har svårt att ge klara besked om vad som händer under huvudvärksepisoderna och om alla symptomens karaktäristika. Därför vore det naturligtvis av stort värde om man kunde bestämma det diagnostiska värdet av enskilda anamnestiska uppgifter vid primära huvudvärksformer, vilket var målsättningen med den aktuella översiktsartikeln.

Naturligtvis frågar man sig vilka svagheter som kan dölja sig i resultaten av en metaanalys av uppgifter som vaskats fram ur publikationer omfattande en tidsperiod på drygt 30 år. Vilken validitet besitter uppgifterna? I de flesta fall är informationen retrospektiv genom att man helt enkelt har samlat in uppgifter från huvudvärkspatienter om egenskaper, symptom, utlösande faktorer med mera utan att närmare försöka kontrollera giltigheten. Man slås av det faktum att i denna typ av undersökningar kommer även en »myt«, förutsatt att den omnämns tillräckligt ofta, att tillmätas lika stort värde som en rad »sanna« observationer. Ett exempel på detta utgör uppgiften om att choklad är en utlösande faktor för migränanfall. Enligt denna översiktsartikel skulle choklad dessutom vara den utlösande faktor som bäst diskriminerar migrän mot huvudvärk av spänningstyp. Detta trots att man i en klinisk kontrollerad undersökning inte kunde bekräfta att choklad är en utlösande faktor för migränanfall hos migräniker som angivit att så skulle vara fallet [6]. Mot denna bakgrund bör man kanske bedöma uppgifterna om de migränutlösande faktorerna med en viss försiktighet. I en dansk epidemiologisk undersökning fann man att huvudvärk av spänningstyp utlöstes av stress, mental spänning och trötthet. Vidare att enbart hos individer med migrän hade episoder med huvudvärk av spänningstyp triggats igång av alkohol, lagrad/övermogen ost, choklad eller fysisk ansträngning [2]. Kanske rörde det sig här i själva verket om »migränanfall« som mer liknade huvudvärk av spänningstyp än migrän?

Författaren har valt att inte utvärdera betydelsen av huvudvärkens intensitet vid respektive huvudvärkstillstånd. Orsaken till detta kan vara att huvudvärkens intensitet inte ingår som ett obligat symptom i den internationella klassifikationen för migrän samt att man i några undersökningar haft svårt att visa att migräniker anger en intensivare och mer smärtsam huvudvärk än de som har huvudvärk av spänningstyp [2]. Emellertid kvarstår det kliniska intrycket att hos individer som besväras av båda huvudvärkstyperna så anges migränanfallen som mer invalidiserande och migränsmärtan som mer intensiv jämfört med huvudvärk av spänningstyp.

Utöver det här redovisade innehållet ur den aktuella översiktsartikeln av Smetana finns i artikeln ytterligare intressant information om migrän, huvudvärk av spänningstyp och Hortons huvudvärk (klusterhuvudvärk) med motsvarande referenser, vilket gör den mycket läsvärd.

Referenser

1. Headache Classification Committee of the International Headache Society. Classification and diagnostic criteria for headache disorders, cranial neuralgias and facial pain. *Cephalalgia* 1988;7:1-96.
2. Ulrich V, Russell MB, Jensen R, Olesen J. A comparison of tension-type headache in migraineurs and in non-migraineurs: a population-based study. *Pain* 1996;67(2-3):501-6.
3. Smetana GW. The diagnostic value of historical features in primary headache syndromes: a comprehensive review. *Arch Intern Med* 2000;160(18):2729-37.
4. Simel DL, Samsa GP, Matchar DB. Likelihood ratios with confidence: sample size estimation for diagnostic test studies. *J Clin Epidemiol* 1991;44(8):763-70.
5. Taube A, Malmquist J. Räkna med vad du tror. Bayes' sats i diagnostiken. *Läkartidningen* 2001;98(24):2910-3.
6. Marcus DA, Scharff L, Turk D, Gourley LM. A double-blind provocative study of chocolate as a trigger of headache. *Cephalalgia* 1997;17(8):855-62

Inbjudan till Läkartidningens läsare och skribenter!

Välkomna till prisutdelningen för bästa artiklar publicerade i LT år 2001.

Möt pristagare, jury samt Läkartidningens redaktion under festliga former på Svenska Läkaresällskapet, Klara Östra Kyrkogata 10, Stockholm, torsdagen den 14 februari.

Program

17.30–18.00
Samling med drink och lätt förtäring.

18.00–19.30
Presentation av tävlingen.
Samtida musik av Hans Persson och Staffan Lang.
Presentation av pristagarna
»Att skriva vetenskap på svenska« – Karin Wahlberg (gynekolog och känd deckarförfattare).

19.30–ca 21.00
Middag (föranmälan).

Föranmälan

(senast den 7 februari)

Till buffén före prisutdelningen till
inga-maj.lagerholm@lakartidningen.se

Till middagen (150 kr) direkt till Läkaresällskapet
gunilla.karlsson@svls.se

Väl mött på Sällskapet den 14:e februari!

Josef Milerad
medicinsk chefredaktör

Marianne Kronberg
chefredaktör

Tony Foucard, docent, överläkare, barnmedicinska kliniken, Akademiska barnsjukhuset, Uppsala

Bengt-Eric Skoogh, professor, överläkare, avdelningen för lungmedicin och allergologi, Sahlgrenska Universitetssjukhuset, Göteborg

Dos–effektkurva för inhalationssteroider vid astma bör beaktas vid dosökning

Kommentar till en Cochrane-rapport

II Inhalationssteroider är i dag grundbulten i all behandling av astma. Beklometasondipropionat (BDP) var den första kortikosteroid som blev tillgänglig för inhalation i början av 1970-talet. Doser på 50–100 µg per dag användes då, i första hand till patienter med svår astma. Med tiden fick man kunskap om att även vid lindrig astma föreligger en inflammatorisk reaktion i bronkialslemhinnan med i typiska fall en dominans av eosinofila celler. Indikationen för inhalationssteroider ökade därmed. Samtidigt har nyare inhalationssteroider blivit tillgängliga, och högre doser än tidigare ges av både BDP och de nyare steroiderna. Eftersom det vid kronisk astma rör sig om långtidsbehandling med inhalationssteroider måste säkerhetsaspekterna noga beaktas. Vid dygnsdoser på 1000 µg eller mer ses en hämning av binjure–hypofysaxeln, och markörer för benomsättning börjar påverkas. För att orsaka kliniskt betydelsefulla biverkningar krävs vanligtvis betydligt högre dygnsdoser. Vilken högsta dygnsdos som kan uppfattas vara utan risk för biverkningar vid långtidsbehandling är svårt att fastställa på grund av biologisk variation, men allmänt anses 200–400 µg per dygn vara en rimligt säker dos [1].

Vid försämring måste effektivare medicinering ges. Man har då att välja mellan att höja dygnsdosen inhalationssteroid eller lägga till ett annat läkemedel, t ex långverkande β₂-agonist eller leukotrienreceptorantagonist, alternativt en kombination av dessa båda åtgärder. Någon bra och gemensam uppfattning om hur mycket dosen inhalationssteroid bör höjas för att få en snabb förbättring till stånd finns inte. Att studera dos–effektsamband vid behandling med inhalationssteroid är därför av praktiskt intresse.

Cochrane-samarbetet har nyligen publicerat en översikt av olika doser BDP vid underhållsbehandling vid astma [2]. Syftet var tvåfaldigt. Dels ville man studera effekt och säkerhet vid behandling med olika doser av BDP, dels se om det fanns ett tydligt samband mellan dos och effekt.

Kriterier för urval av studier

Urvalskriterier var studier på barn över 2 år och vuxna med astma. Om både patienter med astma och kroniskt obstruktiv luftvägssjukdom (KOL) ingick i en studie inkluderades studien om data för astmapatienterna redovisades separat. Be-

SAMMANFATTAT

15 studier som belyser dos–effektförhållandet av beklometason (BDP) vid behandling av astma inkluderades av Cochrane-samarbetet i en metaanalys.

Studierna omfattade behandlingsdoser mellan 100 och 2000 µg per dygn, men få studier belyste extremdoserna.

Tvåfaldig dosskillnad gav i nio studier endast undantagsvis signifikanta förbättringar och endast i enstaka mätvariabler, t ex FEV₁, morgon-PEF, symtom.

Fyr–femfaldig dosskillnad i två studier gav i endast en av dem en förbättring av FEV₁ med 16 procent mot 7 procent med den lägre dosen.

Endast två studier omfattade barn. I dessa studerades tvåfaldig dosskillnad. Endast i den ena kunde en liten förbättring i PEF-värden noteras med den högre dosen.

Dos–effektkurvan är mycket flack, och fyrfaldig snarare än som hittills dubbling av dosen krävs vid otillräcklig effekt. Höga doser under längre tid bör dock undvikas på grund av biverkningsrisk, och andra behandlingsalternativ bör övervägas i dessa situationer.

Evidensbaserad medicin

handlingsperioderna måste vara minst en vecka långa, och det spelade ingen roll om inhalationssteroiden togs som direktverkande spray, spray via andningsbehållare eller som pulver. Som utfallsvariabler användes

- mått på luftvägskaliber (forcerad expiratorisk volym på 1

de senaste åren granskades för att fånga upp studier som ännu inte hunnit publiceras.

Bedömningsförfarande

Två personer bedömde var och en samtliga utvalda arbeten utan att känna till författarnas namn, institution eller vilka som sponsrat arbetet. Varje arbete poängsattes på en skala 0–5 enligt Jadad 1996 [3]. Denna skala innehåller följande frågor:

- Var studien beskriven som randomiserad? (ja=1, nej=0).
- Var studien beskriven som dubbelblind? (ja=1, nej=0).
- Fanns beskrivning av dem som avbröt eller uteslöts ur studien? (ja=1, nej=0).
- Var metoden för randomisering välbeskriven och lämplig? (ja=1, nej=0).
- Var metoden för dubbelblindning välbeskriven och lämplig? (ja=1, nej=0).
- Minska med ett poäng om någon av metoderna för randomisering eller dubbelblindning inte var lämplig.

Författarna tillskrevs åtminstone två gånger om något var oklart med randomisering eller utfallsvariabler. Vid behov kontaktades även tillverkaren för komplettering av data.

Resultat

Femton studier uppfyllde inklusionskriterierna, två studier på barn och 13 på vuxna. Fjorton av studierna fångades upp den normala sökvägen och endast en studie vid granskning av sammanfattning från ett möte. Endast en av studierna gjordes i primärvård, övriga på sjukhus. Sex studier gjordes med parallella grupper, övriga nio var överkorsningsstudier. Resultat från studier med parallella grupper och i överkorsningsstudier har inte slagits samman i redovisningen. Studielängden var en till fyra veckor i nio studier, en till sex månader i tre och mellan sex och tolv månader i två studier.

Studiernas kvalitet varierade. Endast en studie erhöll 5 poäng på Jadad-skalan, tio arbeten erhöll 3–4 poäng medan fyra studier fick 2 eller färre poäng.

Mindre än tvåfaldig skillnad i BDP-dos (400 vs 500 µg/dag). I två överkorsningsstudier (Jadad-poäng 4 respektive 2) i fyra veckor men utan »wash out«-period mellan aktiv behandling och placebo erhöles ingen skillnad i utfallsvariablerna, som var FEV₁, forcerad ventilatorisk kapacitet (FVC), PEF morgon och kväll och i en av studierna dessutom symtomdagbok [4, 5].

Tvåfaldig skillnad i BDP-dos. I en sexveckorsstudie med parallella grupper av patienter som inte fått inhalationskortison tidigare jämfördes doserna 100 och 200 µg/dag med morgon-PEF som utfallsvariabel. Ingen signifikant skillnad erhöles mellan grupperna [6].

I en fyraveckors överkorsningsstudie utan »wash out«-period mellan behandlingsperioderna jämfördes doserna 300 och 600 µg/dygn [7]. Ingen skillnad erhöles mellan grupperna för FEV₁, FVC, PEF morgon och kväll, behov av akut β₂-agonist eller plasmakortisolkoncentration på morgonen.

Doserna 400 och 800 µg/dygn jämfördes i fem studier. I två studier med parallella grupper, den ena på 177 barn i tolv månader [8] och den andra på 365 vuxna i sex veckor [9] (Jadad-poäng 3–4) erhöles en liten men signifikant bättre effekt med den högre dosen BDP i morgon-PEF och en liknande tendens för kvälls-PEF när resultaten slogs samman. Ingen skillnad sågs mellan grupperna i barnstudien för FEV₁, metakolinreaktivitet eller antal som fick avbryta på grund av försämring. I vuxenstudien sågs en liten behandlingsvinst med den större dosen vid jämförelsen i FEV₁, minskning av poäng för

Inhalationssteroider är idag grunden i all behandling av astma.

- sekund (FEV₁), luftflödesvärden (PEF) morgon och kväll, dygnsvariation av PEF enligt dagbok samt PEF-värde vid läkarbesök),
- symptompoäng,
- behov av akut β₂-agonist,
- bronkiell hyperreaktivitet mätt med metakolin eller histamin,
- hälsotillstånd/livskvalitetsformulär,
- astmaförsämring: inläggning på sjukhus, hemma från arbete eller skola, akuta läkarbesök,
- säkerhetsvariabler: serumkortisol, urinkortisol,
- biverkningar i orofarynx: heshet, torrhet, halsont, candidaväxt.

De sökord som användes var glukokortikoid, kortikosteroid, beklometason, budesonid, flutikason, triamcinolon, flunisolid, Becotide, Beclofort, Pulmicort, Flixotide. Alla rapporter som innehöll dessa sökord (N=6 494) fördes till ett särskilt register för inhalationssteroider, ett register som manuellt granskades av två personer för att utesluta dubbelpubliceringar.

I detta register identifierades 1 149 bidrag som föll ut på sökorden beklometason, Becotide, Becloforte. Alla rapporter som inte rörde behandling av kronisk astma hos människa, inte var randomiserade kontrollstudier och inte innehöll en behandlingsarm med inhalationskortison uteslöts. För att inte missa några studier kontaktades tillverkaren av Becotide, GlaxoSmithKline, för att erhålla detaljer om sponsrade studier, och publicerade sammanfattningar från större kongresser

nattliga symtom och minskning av akut behov av β_2 -agonist dagtid.

I två överkorsningsstudier slogs resultaten samman, men inga skillnader mellan doserna kunde ses i utfallsvariablerna FEV₁, FVC, morgon- och kvälls-PEF, behov av akut β_2 -agonist, symtompoäng dag och natt eller procent symtomfria dagar och nätter [10, 11]. I den ena studien undersöktes koncentrationen plasmakortisol före och efter ACTH utan att några skillnader kunde ses mellan grupperna. I en tredje överkorsningsstudie jämfördes under åtta veckor 400 och 800 μg BDP/dygn. Några hade också en liten daglig dos oralt prednisolon, men denna dos hölls oförändrad under studien. Inga signifikant skillnad sågs mellan de båda grupperna avseende morgon- eller kvälls-PEF, FEV₁ eller FVC.

I två små överkorsningsstudier med låg kvalitet (2 Jadad-poäng) jämfördes doserna 500 och 1 000 $\mu\text{g}/\text{dag}$ [12, 13]. I den ena ingick tio patienter som tidigare inte fått inhalationskortison. Efter tre veckors behandling uppvisade de en signifikant förbättring av histaminreaktiviteten med den högre dosen, men inga skillnader sågs för FEV₁, symtompoäng eller behov av akut β_2 -agonist. I den andra studien (N=16) fanns även en behandlingsarm med dygnsdosen 2 000 μg . Inga signifikanta skillnader erhöles mellan behandlingsgrupperna avseende FEV₁, morgon- och kvälls-PEF, dysnépoäng eller behov av akut β_2 -agonist. Inte heller sågs någon skillnad i morgonvärdet av plasmakortisol.

Fyrfaldig skillnad i BDP-dos (400 vs 1600 $\mu\text{g}/\text{dag}$). I en enda studie med parallella grupper (N=423) som behandlades under fyra veckor, studerades denna dosskillnad på patienter som en tid erhållit regelbunden behandling med inhalationssteroid [14]. Studien hade hygglig kvalitet (3 Jadad-poäng). FEV₁ förbättrades med 16 procent mot 7 procent med den lägre dosen ($P<0,03$). Inga signifikanta skillnader sågs däremot för FEV₂₅₋₇₅, FVC, morgon- och kvälls-PEF, symtompoäng, behov av akut β_2 -agonist eller antal som fick avbryta på grund av försämring.

Femfaldig skillnad i BDP-dos (400 vs 2000 $\mu\text{g}/\text{dag}$). I en högkvalitativ överkorsningsstudie (4 Jadad-poäng) på nio vuxna med isocyanatastma med två enveckors behandlingsperioder sågs ingen skillnad mellan grupperna i FEV₁, metakolinreaktivitet eller morgonvärde plasmakortisol [15].

Oral kortisonbehandling. I två studier med hygglig kvalitet (3–4 Jadad-poäng) av parallella grupper under sex månader studerades möjligheten att minska dosen oralt kortison vid behandling med olika doser inhalerad BDP. I en studie randomiserades 148 patienter till 300 eller 1 500 μg BDP/dag [16]. I inledningsfasen av studien hade prednisolondosen anpassats till optimal astmakontroll. Under BDP-behandling sänktes sedan prednisolondosen så mycket som möjligt med i möjligaste utsträckning oförändrade symtompoäng, behov av akut β_2 -agonist och PEF-värden. Ingen skillnad sågs mellan grupperna i dossänkning räknat som mg/dag eller antal patienter som kunde reducera prednisolondosen.

I den andra studien randomiserades 40 patienter till 800 eller 2 000 μg BDP/dag [17]. Även i denna studie sänktes den orala prednisolondosen så mycket som möjligt före randomiseringen till BDP-behandling. Inte heller i denna studie blev det någon skillnad i ytterligare reduktion av prednisolondosen mellan grupperna med medelhög respektive hög dos BDP.

I en tredje studie av överkorsningstyp under åtta veckor jämfördes BDP 400 $\mu\text{g}/\text{dag}$ via spray med BDP 800 $\mu\text{g}/\text{dag}$ som pulver via Rotahaler [18]. En del av patienterna använde dessutom oralt prednisolon, men dosen hölls konstant under hela studien. Ingen signifikant skillnad sågs mellan behand-

lingsgrupperna för någon av utfallsvariablerna morgon- och kvälls-PEF, FEV₁ och FVC.

Författarnas synpunkter

En tidigare systematisk sammanställning har visat att BDP har en signifikant bättre effekt än placebo vid behandling av astma mätt som förbättring i FEV₁, PEF, symtompoäng och behov av akut β_2 -agonist [19]. Däremot kunde inte någon ytterligare effekt av dygnsdoser över 400 μg ses i den sammanställningen. Dessutom finns flera studier som visar att med hjälp av inhalationssteroider kan man reducera dosen oralt prednisolon och i vissa fall övergå till enbart inhalationsbehandling [1]. I den aktuella genomgången har flertalet studier varit överkorsningsstudier med ganska få patienter. För de flesta mätvariablerna har studierna inte uppvisat några signifikanta skillnader. Eftersom använda doser har varierat har resultaten inte kunnat slås samman, vilket gjort det svårt att dra några slutsatser angående eventuellt dos-effektsamband. Däremot har enligt författarna ett visst stöd för en svag dos-effekt kunnat utläsas i studierna med parallella grupper. I jämförelsen mellan 400 och 800 μg BDP/dag kunde man vid sammanslagning av en barn- och en vuxenstudie finna en något bättre effekt med den högre dosen för morgon- och kvälls-PEF [8, 9]. I vuxenstudien gav dessutom den högre dosen något bättre FEV₁, minskade nattsymtom och minskat behov av akut β_2 -agonist [9]. Än större skillnader i FEV₁ sågs i studien som jämförde 400 och 1 600 $\mu\text{g}/\text{dag}$, men inga signifikanta skillnader sågs för övriga mätvariabler [14]. Om dessa tämligen blygsamma skillnader i lungfunktion är kliniskt betydelsefulla är nog tvivelaktigt. Några hållpunkter som talar för att dosen oralt prednisolon kan minskas mer med hjälp av hög än av måttlig dos BDP framkom inte. Författarna avslutar med att påtala behovet av större och bättre studier för att belysa dos-effektsamband vid behandling med BDP och andra inhalationssteroider. I dessa studier bör även utfallsvariablerna innefatta livskvalitet, frånvaro från arbete respektive skola, läkarbesök och sjukhusvård. För att dessa mätvariabler ska bli meningsfulla krävs att studierna är minst sex månader långa och innefattar ett stort antal patienter.

Egna slutsatser

Denna och andra Cochrane-rapporter visar hur svårt det är att göra bra läkemedelsstudier. Ofta slutar rapporterna med att underlaget är kvalitativt och kvantitativt för bristfälligt för att säkra slutsatser ska kunna dras. Dessa rapporter är ändå viktiga, och de kommer sannolikt att leda till att kvaliteten hos läkemedelsstudier successivt förbättras.

Inte oväntat finns det ett dos-effektförhållande även för inhalationssteroider, men relationen är flack och har inte kunnat visas för alla effektvariabler. Den flacka dos-effektkurvan är inte unik för BDP. Samma förhållande tycks gälla även för andra inhalationssteroider som budesonid och flutikason [20–23]. I regel krävs en fyrdubbling av dosen för att i även relativt stora kontrollerade studier kunna påvisa en signifikant effektskillnad [24]. Dessutom tycks dos-biverkningskurvan vara brantare än dos-effektkurvan när väl tröskeldosen för biverkningar nås [24, 25]. Det finns därför goda skäl att vid långvarig underhållsbehandling av astma med inhalationssteroider begränsa användningen av höga doser, vilket även betonas i SBU-rapporten [1]. En lämplig målsättning är att hålla underhållsdosen under 400–800 $\mu\text{g}/\text{dag}$ för BDP och budesonid och under 250–500 $\mu\text{g}/\text{dag}$ för flutikason, där den lägre siffran gäller för barn. Man håller sig då under tröskeldosen för kliniskt märkbara systemeffekter. Den flacka dos-effektkurvan talar dessutom för att man vid otillräcklig effekt av lågdos inhalationssteroid snarare bör överväga komplettering med långverkande β_2 -agonist

och/eller leukotrienreceptorhämmare än att höja dosen inhalationssteroid.

Den genomförda metaanalysen bör tolkas med viss försiktighet. Den flacka dos-effektkurvan är t ex endast visad på populationsbasis. Den behöver inte nödvändigtvis gälla för enskilda astmasjuka. Populationens flacka dos-effektkurva kan betingas av att tröskeldosen för de ingående individerna varierar inom vida gränser. Då blir genomsnittskurvan flack även om individernas dos-effektkurva är brantare. Information saknas på denna punkt.

En annan svaghet är att de flesta studierna hade korta behandlingsperioder. Endast 3/15 studier hade behandlingsperioder på ≥ 3 månader vilket är den tidsgräns som man i SBU-rapporten [1] använt för utvärdering av underhållsbehandling vid astma. Korta behandlingsperioder minskar möjligheten att upptäcka behandlingseffekt på t ex astma-exacerbationer eller bronkiell hyperaktivitet. I en välgjord ettårig studie av behandling med budesonid visades att en fyrdubbling av dagsdosen från 200 till 800 μg gav signifikant färre astmaexacerbationer [23]. Det är också känt att det tar lång tid att uppnå gynnsam effekt på bronkiell hyperaktivitet [26].

Det finns också anledning att påminna om att alla studier av dos-effektsambandet för inhalationssteroider gjorts på astmapatienter i stabilt skede. Resultaten kan därför inte utan vidare appliceras på behandling av akut astmaförsämring. Det är inte säkert att dos-effektsambandet är detsamma vid akut astmaförsämring. I avvaktan på att vi får sådan information känns det rimligt att vid astmaförsämring föreslå en fyrdubbling av inhalationssteroiddosen tills normalisering skett eller högst någon vecka. En sådan strategi stöds av några studier som visat att höga doser inhalationssteroider kan vara lika effektiva som perorala steroider (för referenser se [1]) vid behandling av akut astmaförsämring, även om bedömningsunderlaget ännu är otillräckligt [27]. Kostnadsaspekter kan tala för att även i fortsättningen ge en kort kur med orala steroider. Vi behöver dock fler jämförande studier även på denna punkt.

Referenser

1. Behandling av astma och KOL. En systematisk kunskapssammanställning. Stockholm: SBU 2000.
2. Adams N, Bestall J, Jones P. Inhaled beclomethasone at different doses for long-term asthma (Cochrane Review). In: The Cochrane Library, Issue 1. Oxford: Update Software, 2001.
3. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Controlled Clinical Trials* 1996;17:1-12.
4. Smith MJ, Hodson ME. Twice daily beclomethasone dipropionate administered with a concentrated aerosol inhaler: efficacy and patient compliance. *Thorax* 1986;41:960-3.
5. Carmichael J, Duncan D, Crompton GK. Beclomethasone dipropionate dry-powder inhalation compared with conventional aerosol in chronic asthma. *BMJ* 1978;2:657-8.
6. Hampel F, Lisberg E, Vanden Burgt J, Henon C, Stampone P. 50 mcg twice daily of ultrafine HFA-beclomethasone dipropionate aerosol improves asthma control in adult patients. *Am J Resp Crit Care* 1977;155:A666.
7. Lal S, Malhotra SM, Gribben MD, Butler AG. Beclomethasone dipropionate aerosol compared with dry powder in the treatment of asthma. *Clin Allergy* 1980;10:259-62.
8. Verberne AA, Frost C, Duivenman EJ, Grol MH, Kerrebijn KF. Addition of salmeterol versus doubling the dose of beclomethasone in children with asthma. *Am J Resp Crit Care Med* 1998;158:213-9.
9. Drepaal BA, Payler DK, Qualtrough JE, Perry LJ, Bryony F, Reeve A, et al. Becotide or becodisks? A controlled study in general practice. *Clinical Trials Journal* 1989;26:335-44.
10. So SY, Lam WK. Twice daily administration of beclomethasone dipropionate dry-powder in the management of chronic asthma. *Asian Pac J Allergy Immunol* 1986;4:129-32.
11. Wolthers OD, Pedersen S. Short term growth during treatment with

inhaled fluticasone propionate and beclomethasone dipropionate. *Arch Dis Child* 1993;68:673-6.

12. Carpentiere G, Marino S, Castello F, Balanza C, Bonanno CT. Dose-related effect of beclomethasone dipropionate on airway responsiveness in asthma. *Respiration* 1990;57:100-3.
13. Molema J, Lammers JW, van Herwaarden CL, Folgering HT. Effects of inhaled beclomethasone dipropionate on beta 2-receptor function in the airways and adrenal responsiveness in bronchial asthma. *Eur J Clin Pharmacol* 1988;34:577-83.
14. Nathan RA, Nalop KB, Cuss FM, Lorber RR. A comparison of double-strength beclomethasone dipropionate (84 mcg) MDI with beclomethasone dipropionate (42 mcg) MDI in the treatment of asthma. *Chest* 1997;112:34-9.
15. De Marzo N, Fabbri LM, Crescioli S, Plebani M, Testi R, Mapp CE. Dose-dependent inhibitory effect of inhaled beclomethasone on late asthmatic reactions and increased responsiveness to methacholine induced by toluene diisocyanate in sensitised subjects. *Pulm Pharmacol* 1988;1:15-20.
16. Hummel S, Lehtonen L. Comparison of oral-steroid sparing by high-dose and low-dose inhaled steroid in maintenance treatment of severe asthma. *Lancet* 1992;340:1483-7.
17. Tarlo SM, Broder I, Davies GM, Leznoff A, Mintz S, Corey PN. Six-month double-blind, controlled trial of high dose, concentrated beclomethasone dipropionate in the treatment of severe chronic asthma. *Chest* 1988;93:998-1002.
18. Chatterjee SS, Butler AG. Beclomethasone dipropionate in asthma: a comparison of two methods of administration. *British Journal of Diseases of the Chest* 1980;74:175-9.
21. Busse WW, Chervinsky P, Condemi J, Lumry WR, Petty TL, Renard S, et al. Budesonide delivery by Turbuhaler is effective in a dose-dependent fashion when used in the treatment of adult patients with chronic asthma. *J Allergy Clin Immunol* 1998;101:457-63.
24. Pedersen S, O'Byrne P. A comparison of the efficacy and safety of inhaled corticosteroids in asthma. *Allergy* 1997;52:1-34.

I Läkartidningens elektroniska arkiv
<http://tarkiv.lakartidningen.se>
är artikeln kompletterad med fullständig referenslista.

SUMMARY

Dose-effect relation with inhaled steroids in asthma should be paid attention to when the dose is increased. Comment on a Cochrane-review

Tony Foucard, Bengt-Eric Skoogh

Läkartidningen 2002; 99:264-7

The Cochrane collaboration has performed a meta-analysis of all studies found on the dose-effect relation with beclomethasone dipropionate (BDP) in the treatment of asthma. Fifteen studies were found and included, of which only two comprised children. Daily doses of 100 to 2000 μg were used. A two-fold increased dose was used in 9 studies but only few significant improvements were seen with the higher dose. A four-to fivefold increased dose gave significant improvement in one of two studies with an improvement of FEV1 of 16% versus 7% with the lower dose. The dose-effect of BDP seems to be very flat and a four-fold increase of the dose seems more reasonable than the two-fold commonly used today, when improved effect is desired. This increases the risk of side effects, and for long-term therapy other alternatives should be considered.

Correspondence: Tony Foucard, Barnmedicinska kliniken, Akademiska barnsjukhuset, SE-751 85 Uppsala.

Cochrane-översikt om att förebygga navelinfektioner

Antiseptisk lösning är onödig i i-länder men livräddande i u-länder

II Varje år avlider omkring 500 000 nyfödda spädbarn av neonatal tetanus, och ytterligare ca 460 000 dör av andra bakteriella infektioner. Årligen dör alltså omkring en miljon nyfödda spädbarn av bakterieinfektioner som i många fall utgår från naveln. Nyfödda med tetanus har ofta en samtidig infektion med andra bakterier i naveln. Detta tyder på gemensamma orsaker till infektionerna, främst bristande hygien vid förlossning och navelvård [1].

Studier från fattiga länder och slumområden har visat på upp till 3 procents incidens av neonatal sepsis [1]. I Sverige är problemet betydligt mindre än i fattiga länder, och neonatal tetanus förekommer inte. År 1998 avled 16 spädbarn i Sverige av »infektioner specifika för den perinatale perioden« enligt Dödsorsaksregistret [2].

I denna artikel skall jag redogöra för en Cochrane-översikt som försöker besvara frågan: Finns det någon form av navelvård som förhindrar infektioner bättre än att bara hålla naveln torr och ren [3]?

En allvarlig infektion

Navelsträngen är en unik vävnad med två artärer och en ven inbäddade i Whartons gel som består av hyaluronsyra och omges av en tunn slemhinna (Figur 1). Någon minut efter barnets framfödande upphör pulsationerna i naveln, navelartärerna dras ihop, och barnmorskan skär av naveln. Naveln får inte längre någon blodtillförsel och under de närmaste dagarna börjar den torka in och bli svart och stel (torrt gangrän). Naveln faller av 5–15 dagar efter födelsen.

En navelinfektion, omfalit, kan hindra navelkärnen från att obliterera. Bakterierna får därigenom tillträde till blodet. I navelkärnen kan det bildas septiska tromber som kan sända septiska embolier till exempelvis lungorna. Infektionen kan också spridas direkt från naveln in till bukhålan och förorsaka en peritonit (Figur 1). Omfalit är sålunda en allvarlig infektion, som skall behandlas aggressivt med antibiotika intravenöst. Omfalit känns igen på sedvanliga inflammationstecken: rodnad, svullnad, ömhet (Figur 2).

Lättillgänglig handdesinfektion minskar förekomsten

Epidemiologiska studier av handtvätt och handdesinficering inom sjukvården har visat att läkare, manligt kön, överbe-

SAMMANFATTAT

En miljon nyfödda barn dör årligen av bakteriella infektioner som ofta kommit in i kroppen via naveln.

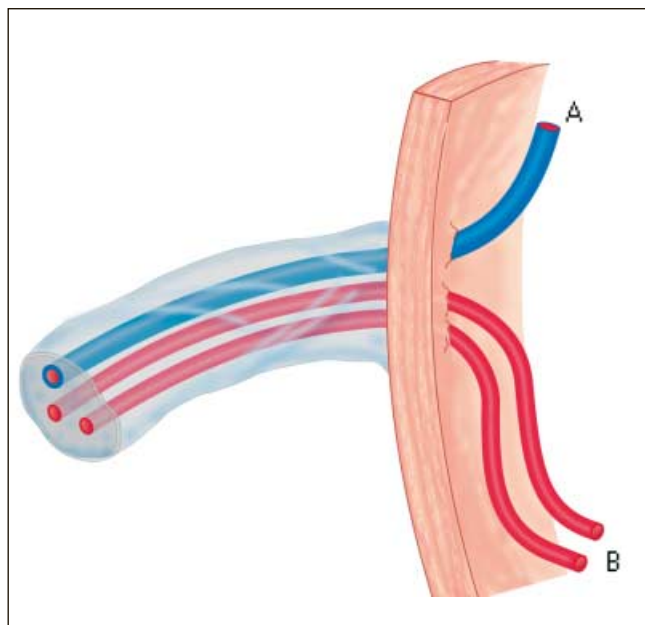
En Cochrane-översikt visar att i rika länder som Sverige är allvarliga navelinfektioner sällsynta. Tvål och vatten bör räcka för navelvård; ingen annan metod har visats vara överlägsen. Det finns dock inte underlag för att bedöma skillnader i incidens av sepsis efter olika navelregimer. På neonatalavdelningar bör man fortsätta att tvätta navlar med klorhexidinlösningar.

I fattiga länder tyder observationsstudier på att om gravida kvinnor vaccinerats mot stelkramp och skadliga traditioner för navelvård ersätts med desinficerande puder eller lösningar kan många spädbarns liv räddas.

Evidensbaserad medicin

läggning och personalbrist är riskfaktorer för slarv. Kvinnor och sjuksköterskor är renligast, manliga läkare slarvigast. Lättillgänglig handdesinfektion, t ex handsprit vid varje kувös och sång på en neonatalavdelning, förbättrar personalens handhygien och minskar förekomsten av nosokomiala infektioner. För en jäktad sköterska kan desinfektion inom armlängds avstånd vara en förutsättning för att bestämmelser om handdesinfektion skall kunna efterlevas. Ibland kräver dessa bestämmelser att sköterska ägnar en fjärdedel av sin arbetstid till att tvätta händerna. Det är också viktigt att tvättspriten innehåller glycerol så att händerna inte blir torra och fulla med sprickor [5].

I Sverige använder vi i huvudsak klorhexidin för desinfektion av navlar och händer. Klorhexidin säljs som vattenlösning i koncentrationer från 0,5 mg/ml (0,05 procent) till 40 mg/ml (4 procent) och som klorhexidinsprit (etanol 70 pro-



Figur 1. Ett tvärsnitt genom bukväggen med navelstump med en ven (a) och två artärer (b) [1].

cent v/v [volymsprocent] = 60 procent w/v [viktpcent]; klorhexidin 5 mg/ml = 0,5 procent).

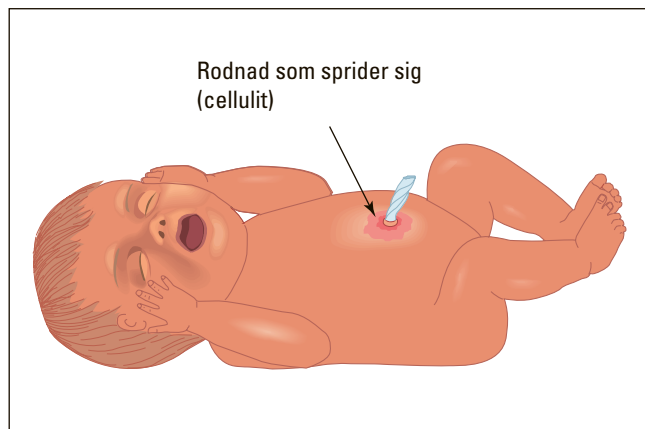
II Material och metoder

Den aktuella systematiska översikten publicerades i Cochrane-bibliotekets »Pregnancy and childbirth group« 1998 [3]. Cochrane-översikter skall efter hand uppdateras, men inga större ändringar är hittills gjorda i denna översikt. Arbetet är en del av ett projekt inom World Health Organization (WHO), som presenteras utförligare i en WHO-rapport från 1999 [1].

Inklusionskriterier för litteratursökningarna var att studierna skulle vara randomiserade eller kvasirandomiserade och jämföra olika metoder för navelvård. Studiernas primära effektmått skulle vara kliniska tecken till navelinfektioner eller till systemisk bakteriell infektion eller död. Kvasirandomisering innebär oftast att varannan patient behandlas med metod 1 och resten med metod 2 (alternerande allokering).

Man sökte i Cochrane Controlled Trials register (sökord: »Umbilical and Cord*«) och i Medline (sökord ej angivna) samt kontaktade WHO och experter inom området för att få tips om fler kliniska studier.

Alla relevanta studier granskades av de två författarna. Kunde de inte enas om huruvida artikeln skulle vara med i undersökningen eller ej hänsköts avgörandet till en tredje person. Tio studier uppfyllde inklusionskriterierna: fyra från USA, två från vardera Kanada och Norge och en från vardera Israel och Storbritannien. Alla undersökningar var begränsade till friska, fullgångna barn födda på sjukhus i i-länder. WHO:s målsättning att få fram undersökningar från fattiga länder och hemförlösningsmisslyckades alltså. De tio studierna jämförde antiseptiska preparat med ingen speciell behandling (fem studier), antiseptiskt puder med adstringerande dito (en studie), antiseptiskt preparat med hydrofob kompress (en studie), antibiotikum med trippelfärg (triple dye; en studie), antibiotikum med ingen särskild behandling (en studie) eller dagligt bad med bara ett initialt bad (en studie). Åtta av studierna var randomiserade. Längst uppföljning var sex veckor, och störst antal patienter (N=2 441 i en undersökning) har Alf Meberg i sina studier från Sentralsygehuset i Vestfold i Tönsberg, Norge [6, 7]. I



Figur 2. »Röd hud runt naveln är ett allvarligt tecken«, skriver King, King och Martodipoero i sin klassiska lärobok i barnhälsovård och fortsätter: »Det är ett farligt tecken, ty det visar att en infektion sprider sig och orsakar cellulit. Barnet kan få sepsis och dö.« [4].

övriga studier har barnen i allmänhet bara följts upp tills naveln ramlat av.

II Resultat

Översiktens resultat sammanfattas i Tabell I och II. Resultaten är magra. Ingen patient drabbades av någon allvarlig, invasiv infektion eller död. Man hittar inga påtagliga fördelar eller nackdelar med någon metod för navelvård. Lokalbehandling med alkohol eller antibiotika får naveln att sitta kvar längre än normalt, vilket oroar en del mammor men tycks i övrigt inte ha för- eller nackdelar. Antiseptika och antibiotika minskar koloniseringen av naveln med *Staphylococcus aureus* m fl bakterier, men det är ovisst om det leder till färre infektioner. Resultaten gäller bara för fullgångna barn i rika länder.

II Diskussion

Författarna konkluderar att deras översikt inte visar vilka metoder för navelvård som är bäst för barnen. Grundregeln att hålla naveln torr och ren gäller fortfarande: ingen annan metod är bevisat bättre för fullgångna, friska barn i höginkomstländer.

Cochrane-författarna fann inga välgjorda undersökningar av spädbarn med hög risk för sepsis: prematurer, barn på intensivvårdsavdelningar och barn i fattiga länder, dvs de flesta barn. Dessa barn löper ökad risk för sepsis. Antiseptika är sannolikt inte skadliga för dem och kan kanske reducera risken för allvarliga infektioner. Författarna drar slutsatsen att man bör fortsätta med antiseptika till dessa barns navlar.

I fattiga länder bör randomiserade, kontrollerade kliniska prövningar genomföras med jämförelser mellan olika regimer för navelvård. En metod som bör testas är användning av kolostrum på naveln. Författarna konkluderar att framtida forskning bl a bör besvara följande frågor: Vilket är naturlöslöppet när navelstumpen faller? Vilka preparat är bäst för att hindra infektioner att tränga in i kroppen genom naveln: alkohol, antiseptika, antibiotika eller kolostrum? Puder, lösningar eller salvor – när skall de användas? Ett lämpligt första projekt i ett låginkomstland vore att erbjuda mödrar utrustning för navelvård och att randomisera grupper av mödrar till tvål och vatten eller klorhexidinlösning.

Svårt att bedöma Cochrane-översiktens tillförlitlighet

Miljontals nyfödda barn behandlas varje år på olika sätt för att förhindra navelinfektioner. Trots detta ger de kontrolle-

Tabell I. Resultat av randomiserade studier med jämförelse av navelvård med respektive utan antiseptika. NNT = number needed to treat (har beräknats från tabelldata i referens 3). NNT = 22 för mödrars oro för naveln betyder att ca 22 navlar måste alkoholbehandlas för att en mor ska lugnas.

Studieobjekt	Resultat	Anmärkningar
Kolonisering av stafylokocker	Färre navlar koloniserade efter vissa antiseptika	
Tid tills naveln fällts	Naveln fälldes 1,6 dagar senare efter alkoholtvätt	P<0,01, N=1 811
Navelinfektion	Ingen skillnad, ingen tydlig trend	P=0,8, N=2 360
Systemisk infektion	Inga systeminfektioner förekom	N=1 811
Död	Inga dödsfall förekom	N=1 811
Mödrars oro över naveln	Färre mödrar var oroliga (NNT=22) och något färre barn fick antibiotika (NNT=117) efter isopropanol på naveln jämfört med ingen behandling	N=1 807

Tabell II. Resultat av randomiserade studier med jämförelse av olika metoder för navelvård.

Jämförda metoder för navelvård	Resultat
Antibiotika/antiseptika	Minskad kolonisation med <i>Staphylococcus aureus</i> efter bacitracin (N=283) men inga fall av infektion i någon grupp (N=205)
Puder/inget puder	Naveln fälldes något tidigare efter puder i en undersökning
Puder en gång/puder varje dag	Ingen skillnad i tid tills naveln fälldes
Kompress/öppen behandling	Inga kliniskt betydelsefulla skillnader
(Salvkompress jämfördes med daglig rengöring med 0,5-procentig klorhexidin i 70-procentig etanol; N=2 441)	
Helkroppstvätt/ej tvätt	Ingen skillnad i kolonisering med <i>Staphylococcus aureus</i>

rade prövningarna av dessa behandlingar inga resultat som kan tillämpas på underburna eller sjuka barn, eller barn i u-land, eller på svåra bakteriella infektioner. De kontrollerade studierna är få och små. Detta överraskar författarna: »Few interventions for cord care have been evaluated by randomized controlled trials. This is surprising given the millions of newborns in developed countries that are treated one way or another.« The Cochrane library började som en databas i perinatalmedicin. Trots detta vilar många metoder i perinatalmedicinen fortfarande på ett vetenskapligt gungfly [8].

Författarna sökte elektroniskt i Cochrane Controlled Trials register (sökord: »Umbilical and Cord*«) och fick 146 träffar 1997. Jag fick 340 träffar när jag gjorde om sökningen i december 2001; alltså mer än dubbelt så många. I juni 2001 sökte jag »Umbilical and Infection« på Medline. De 200 första träffarna (arbeten publicerade ungefär oktober 1999–juni 2001) verkade inte innehålla någon betydande nyhet om navelvård och navelinfektioner.

I den WHO-rapport [1] som delvis sägs bygga på Cochrane-översikten refereras en studie från Pakistan som visat att lokalbehandling av navelstumpen med antibiotika effektivt förebygger tetanus [9]. Denna undersökning, som är en fall-kontrollstudie, bör diskuteras i en uppdatering av Cochrane-översikten. Visserligen uppfylls inte inklusionskriterierna om randomiserad studie, men när man inte lyckats hitta en enda randomiserad studie från u-länder bör man granska denna välgjorda observationsstudie.

Det finns en osäkerhet i författarnas rekommendationer eftersom ingen av studierna innehöll några sepsisfall. Detta visar att neonatal sepsis är en sällsynthet bland fullgångna, väl-skapta barn i rika länder. Även en liten skillnad i risken för sepsis kan emellertid vara kliniskt signifikant eftersom risken för handikapp eller död är stor. Undersökningar från Sverige [10] och Norge [6] [Alf Meberg, Tønsberg, pers medd, 2001] visar att incidensen av ytliga infektioner mer än halverades

när man gick över från tvål och vatten till Hibiscrub (4-procentig klorhexidin-/tvållösning) vid navelvård på BB. Det är därför tänkbart att även förekomsten av sepsis är olika med och utan klorhexidin. Å andra sidan kan nutidens rooming-in och korta vårdtider på BB ha minskat risken för sepsis hos barnen oavsett typ av navelvård.

Jag tycker sammanfattningsvis att Cochrane-översikten bör uppdateras eftersom nya artiklar tillkommit, observationsstudier (kohort- och fall-kontrollstudier) från u-länder inkluderas, sökvägarna anges betydligt bättre, och hänsyn måste tas till att sepsisfall saknas i undersökningarna. Det är nu svårt att bedöma hur pålitlig översikten är.

Banbrytande upptäckt

Det känns inte rätt att diskutera hygien på förlossnings- och BB-avdelningar utan att nämna Ignaz Philipp Semmelweis (1818–1865). Semmelweis föddes i Budapest, utbildade sig till läkare i Wien och till gynekolog vid Allgemeines Krankenhaus i Wien. Till dess stora förlossningsklinik fördes ogifta, gravida kvinnor och andra kvinnor från samhällets skuggsida. Barnsängsfeber skördade skrämmande antal dödsfall bland nyförlösta mödrar och nyfödda barn på Första kliniken, där läkare utbildades. Andra kliniken, som låg vägg i vägg med Första kliniken och där barmorskor utbildades, hade betydligt lägre dödlighet. När Semmelweis och hans assistenter i maj 1847 började tvätta händerna i klorinhaltigt vatten efter morgonens obduktioner sjönk mortaliteten bland patienterna på läkarkliniken snabbt ned till barmorskeklinikens nivå. Barmorskorna ägnade sig aldrig åt obduktioner.

Semmelweis insåg att han gjort en banbrytande upptäckt: mödrarna och de nyfödda barnen dödades av ett smittämne som överfördes från en död mor till en ännu levande via läkarnas händer. Hans klassiska bok »Die Aetiologie, der Begrifff, und die Prophylaxis des Kindbettfiebers« (1860) [11] utgör ännu efter 140 år en faktiskt mycket lärorik läsning för

epidemiologiskt intresserade. Nu vet vi att barnsängsfeber orsakas av streptokocker. Semmelweis' noggranna handhygien är lika viktig idag som den var 1847.

När barnsalar infördes på BB under 1940-talet blev de nyfödda barnens navlar i ökad utsträckning koloniserade med *Staphylococcus aureus*, vilket beredde vägen för stafylokokkinfektioner i hud och navelstump. Under 1980- och 1990-talen infördes rooming-in, stöd av amning och hudkontakt mellan barn och föräldrar och begreppet »baby-friendly hospital« [12], vilket minskar risken för smitta mellan barnen [1].

Mina konklusioner

Enligt Zupan och Garner bör vi i Sverige och andra rika länder hålla barnens navlar torra och rena på BB med hjälp av tvål och vatten. Klorhexidin bör reserveras för barnen på neonatalavdelningar. Jag kan med viss tvekan och reservation acceptera deras slutsatser. Jag anser att avsaknaden av sepsisfall medför att slutsatserna bara kan betraktas som preliminära och att vi därför bör följa upp incidensen av neonatal sepsis vid olika regimer för navelvård. Vi har här »no evidence of effect« av klorhexidin, men vi har inte »evidence of no effect«. Författarna gör det vanliga felet att inte hålla isär dessa två begrepp [13].

I Sverige kan vi för en uppföljning använda våra perinatale databaser. De genomgår nu en förändring som förhoppningsvis innebär en mer omfattande datainsamling om perinatal hälsa och sjuklighet. En rundringning till en handfull svenska BB-avdelningar visar att svenska navelstumpar tvätas med klorhexidinsprit eller vattenlösning av klorhexidin eller bara tvål och vatten, beroende på vilket BB barnen vårdas på. Det går nog inte att nu randomisera svenska BB till olika rutiner, men det bör finnas möjlighet till en observationsstudie för att utröna om olika navelregimer innebär olika risk för sepsis.

I u-länder med förekomst av neonatal tetanus bör gravida kvinnor vaccineras mot tetanus (antikroppar går då över i fostrets blod) och lära sig navelhygien med desinficerande/antimikrobiella lösningar eller puder i stället för kodynga. Jordemödrar (traditional birth attendants) skall lära sig ren förlossning (clean delivery).

Cochrane-översikten bör uppdateras och förbättras (se ovan). Cochrane-översikter skall vara öppna för »post-publication peer review«. Jag ämnar därför sända mina kommentarer till Cochrane »Pregnancy and Childbirth Group« på det sätt som beskrivits i Läkartidningen [14].

Referenser

1. WHO. Care of the umbilical cord: a review of the evidence. Geneva: World Health Organization; 1999. www.who.int/rht/documents/MSM98-4/MSM-98-4.htm
2. Epidemiologiskt centrum. Dödsorsaker 1998. Statistik: Hälsa och sjukdomar 2001:1. Stockholm: Socialstyrelsen; 2001. www.sos.se
3. Zupan J, Garner P. Topical umbilical cord care at birth (Cochrane Review). In: The Cochrane Library, Issue 1. Oxford: Update Software; 2002.
4. King M, King F, Martodipoero S. Primary child care: book one. A manual for health workers. Oxford: Oxford University Press; 1978.
5. Pittet D, Boyce JM. Hand hygiene and patient care: pursuing the Semmelweis legacy. *Lancet Infectious Diseases* 2001;0:9
6. Meberg A, Schøyen R. Bacterial colonization and neonatal infections. Effects of skin and umbilical disinfection in the nursery. *Acta Paediatr Scand* 1985;74:366-71.
7. Meberg A, Schøyen R. Hydrophobic material in routine umbilical cord care and prevention of infections in newborn infants. *Scand J Infect Dis* 1990;22:729-33.
8. Silverman WA. Where's the evidence? Debates in modern medicine. Oxford: Oxford university press; 1998.
9. Bennett J, Macia J, Traverso H, Banoagha S, Malooly C, Boring J. Protective effects of topical antimicrobials against neonatal tetanus. *Int J Epidemiol* 1997;26:897-903.
10. Seeberg S, Brinkhoff B. Epidemiology and control of staphylococcal pyoderma among newborn infants: evaluation of a method for routine cord care with 4 per cent chlorhexidine-detergent solution. *J Hosp Infect* 1984;5:121-36.
11. Semmelweis I. Die Aetiologie, der Begriff, und die Prophylaxis des Kindbettfiebers. Pest, Wien, Leipzig: CA Hartleben's Verlag; 1861. Förkortad, kommenterad, amerikansk upplaga: The etiology, concept, and prophylaxis of childbed fever. Madison, Wisconsin: The University of Wisconsin Press; 1983.
12. Unicef. BFHI news. The baby-friendly hospital initiative Newsletter. Oct/Dec 2001. New York. www.unicef.org/bfhi/
13. Tarnow-Mordi WO, Healy MJ. Distinguishing between »no evidence of effect« and »evidence of no effect« in randomised controlled trials and other comparisons. *Arch Dis Child* 1999;80:210-1.
14. Pildal J, Goetzsche PC. Kommentar: Kritikken er berettiget. *Läkartidningen* 2001;98:5625.

SUMMARY

Topical umbilical cord care at birth

Inge Axelsson

Läkartidningen 2002;99:1563-6

One million newborn infants die every year by bacterial infections, which often have entered the body via the umbilicus. A Cochrane systematic review on »Topical umbilical cord care at birth« by J Zupan and P Garner is reviewed. Zupan and Garner conclude that simply to keep the umbilical cord dry and clean is sufficient for healthy, term neonates in rich countries; disinfectants do not offer any advantage. However, cleaning the umbilical cord with disinfectants may reduce the risk of serious bacterial infections in babies in poor countries or in neonatal wards. Observational studies in poor countries indicate that the life of numerous infants can be saved if pregnant women are vaccinated against tetanus and disinfectants are substituted for harmful cord care traditions.

This Cochrane review is credible, but it should be updated and considered tentative since no data on sepsis are included. The search strategy should be better described and observational studies (case control and cohort studies) from poor countries should be included since there are no randomized control trials from these countries.

Correspondence: Inge Axelsson, Dept of Pediatrics, Östersunds sjukhus, SE-831 83 Östersund, Sweden.

Jon Karlsson, professor, överläkare, ortoped, Sahlgrenska Universitetssjukhuset/Östra, Göteborg
(jon.karlsson@vgregion.se)

Stöd för fotleden förebygger ledbandsskador

■ Supinationsskador, särskilt lateralt i fotleden, som leder till ledbandsruptur av varierande grad är bland de vanligaste skadorna både inom primärvård och på akutmottagningar världen över. Speciellt utsatta är unga aktiva idrottsutövare. Ungefär 5 000 sådana skador behandlas dagligen, huvudsakligen i primärvården i England [1]. Hos människor med inaktiv livsstil är skadan smärtsam, men den leder sällan till vare sig frånvaro från arbete eller allvarliga resttillstånd. Hos både idrottsmän och individer med krävande arbete kan en ledbandsskada ha stor betydelse både på kort och på lång sikt.

De flesta av dessa skador läker dock med rehabilitering, utan operativ behandling. Det finns ett stort antal behandlingsstudier som visar goda resultat hos 80–90 procent av patienterna.

Skadorna är däremot mycket vanliga i idrottsgrenar såsom fotboll, handboll, basketboll och volleyboll. Den vanligaste skademekanismen är samtidig inåtrotation och plantarflexion. Det främre ledbandet, ligamentum talo-fibulare anterior, är det som alltid skadas först. Därefter följer varierande grad av skada på kalkaneofibulära ligamentet. Det bakre ligamentet, ligamentum talo-fibulare posterior, skadas mycket sällan. Tillsammans utgör dessa tre ligament det laterala ligamentkomplexet. Även om isolerade mediala (ligamentum deltoideum) ledbandsskador förekommer överväger laterala ledbandsskador stort i antal.

Skadeindelning

Traditionellt indelas dessa skador i Grad I–III. Vid en klinisk undersökning är det dock svårt att skilja mellan de olika graderna; denna distinktion har sannolikt liten betydelse för behandlingsresultat och eventuella resttillstånd. Teoretiskt är en Grad I-skada lindrig och representerar uttöjning av ledbanden, Grad II är medelsvår och utgör en partiell ruptur av ledbanden, medan Grad III är allvarlig och utgör total ruptur [1].

De flesta skador läker av sig själva, men ibland förekommer mer eller mindre allvarliga resttillstånd, t ex ökad laxitet (förlängning av ledbanden), nedsatt muskelkraft i peroneusmuskulaturen och nedsatt balansförmåga. Detta leder till kronisk funktionell instabilitet, dvs att individen trampar snett ofta utan egentlig anledning. Det föreligger ingen säker korrelation mellan laxitet och funktionell instabilitet.

Eftersom ledbandsskador i fotleden är så vanliga (flera in-

SAMMANFATTAT

En systematisk översikt sammanställd inom Cochranesamarbetet refereras, där olika typer av intervention för att förebygga ledbandsskador i fotleden har studerats.

Särskilt hos individer som tidigare har ådragit sig ledbandsskada kan risken för skada minskas genom användning av ortos, som ger en signifikant reduktion av antalet fotledsstukningar hos utövare av högriskidrotter som fotboll, basketboll och volleyboll.

Det är bättre att använda någon form av yttre stöd än inget alls.

Viss vetenskaplig information talar för en förebyggande effekt av balansträning.

Evidensbaserad medicin

cidensstudier visar att en ledbandsskada utgör 15–45 procent av alla idrottsrelaterade skador) finns det ett stort intresse för förebyggande åtgärder, framför allt bland idrottsutövare. De metoder som har testats är bl a anpassning av skor, användning av olika skydd såsom ortoser och idrottstejp samt specifika träningsprogram och träningsråd. Se k proprioceptiv träning med hjälp av balansbräda används ofta i förebyggande syfte, mestadels på empiriska grunder.

Det stora antalet skador orsakar också sammantaget stora kostnader, även om varje enskild skada inte leder till så lång bortovaro från idrott eller arbete. En svensk studie har visat en ekonomisk besparing på inte mindre än 60 miljoner kronor årligen vid förbättrad behandling av laterala ledbandsskador i fotleden.

Hypoteser och kriterier

Syftet med föreliggande artikel är att referera en systematisk översikt sammanställd inom Cochranesamarbetet av



Figur 1.

Ett yttre fotledsstöd kan ha god förebyggande effekt mot ledbandsskador. Det är försett med luftkuddar på insidan, s k »air-cast«, och består av ett yttre plasthölje som hålls ihop med ett hälband och ett kardborreband runt fotleden.



Quinn och medarbetare [2], där olika typer av intervention för att förebygga ledbandsskador hos individer från tonår till medelålder redovisas. Evidensbaserade fakta visar att det är möjligt att förebygga ledbandsskador i fotleden, en banal men mycket kostsam skada. Förutom randomiserade och semirandomiserade studier baseras rapporten även på en metaanalys. I de fall det var möjligt analyserades individer utan tidigare skador separat.

De nollhypoteser som testades var:

- Ingen skillnad förekommer vad gäller intervention som syftar till att förebygga ledbandsskador och ingen intervention alls.
- Ingen skillnad föreligger vad gäller olika metoder som syftar till att förebygga ledbandsskador.

Kriterier för att inkludera studier var:

- Alla randomiserade eller semirandomiserade kliniska prövningar angående förebyggande interventioner för ledbandsskador i fotleden.
- Individer från tonårsålder till medelålder med risk för eller som tidigare hade ådragit sig ledbandsskador inkluderades i studien.
- Kliniska studier där någon intervention som syftade till att minska risken för ledbandsskador testades. De olika interventionerna inkluderade t ex justering av skor, användning av yttre stöd såsom ortos alternativt idrottstejp, olika träningsprogram i specifikt syfte att förebygga ledbandsskador. Studier som involverade enbart förebyggande åtgärder i laboratoriemiljö utan direkt klinisk relevans uteslöts.
- Som effektmått evaluerades prevention med hänsyn till in-

cidens av ledbandsskada, allvarlighetsgrad av ledbandsskada (Grad I–III, behov av operation), komplikationer (funktionsnedsättning, hudskador, andra skador), utvärdering av resursanvändning (antal besök vid sjukvårdsinrättning, kostnad) samt slutligen subjektiv utvärdering av funktionell instabilitet, »giving way«, och nedsatt funktion.

Databassökning

Sökning genom användning av databaserna Medline, Embase, Cinahl och Current Contents, Cochranes eget register över Controlled Clinical Trials (CCTR), gjordes. Nivå 1 och nivå 2 av Cochrane Optimal Strategy, 26 sökord, alternativt kombinationer, användes.

II Metod

Studier inkluderades och utvärderades med hänsyn till metodologisk kvalitet. Studierna bedömdes enligt Cochranekriterierna och poängsattes med poängen 3, 2 eller 1. Följande frågor ställdes:

- Är beskrivning och genomförande av randomiseringen adekvat?
- Beskrevs patienter som utgick ur studien men inkluderades i analysen, dvs »intention to treat«?
- Var studien blindad med hänsyn till behandlingsstatus?
- Var behandlingsgrupp och kontrollgrupp jämförbara vid inklusion?
- Var de individer som inkluderades blindade med hänsyn till status efter inklusion?
- Var utvärderade blindade med hänsyn till behandling?
- Var behandlingsprogrammen annat än vad gäller den primära behandlingsregimen identiska?
- Var inklusions- och exklusionskriterierna klart definierade?
- Var slutpunkt med hänsyn till slutresultat klart definierad?
- Var de diagnostiska test som användes kliniskt användbara?
- Förekom hänvisning till t ex radiologiska test, magnetkameraundersökning, undersökning av s k »talar tilt« (draglåda) samt beskrivning av speciell klinisk behandlingsalgoritm?
- Var behandlingstiden kliniskt relevant?

Studierna poängsattes med maximal poängsumma 33.

Studierna

Elva randomiserade studier identifierades från 1966 till 2001. Fem studier [3-7], som inkluderade sammanlagt 3 954 individer, följde inklusionskriterierna enligt ovan. Ytterligare fem studier exkluderades eftersom de inte innehöll information om förebyggande åtgärder [8-10] eller var laboratorie-

studier fokuserade på biomekanik och postural kontroll [11, 12]. Vissa av studiernas effektmått handlade om den nedsatta förmåga till idrottsaktivitet som följer med användning av ortos. Den elfte studien [13], som testade användning av höga eller låga skor kompletterade med idrottstejp eller ingen tejp alls, hade alltför bristfällig uppföljning för att kunna inkluderas. Samtliga studier involverade aktiva unga individer som deltog i organiserade aktiviteter. Endast två studier [3, 4] inkluderade kvinnor (8,3 procent respektive 4 procent). I de studier där rapporter förelåg var medelåldern mellan 19 och 23 år.

Andelen deltagare med anamnes om tidigare ledbands-skador varierade från ungefär 10 procent [5, 7] till mer än 50 procent [6], eller nämnades inte. Samtliga studier genomfördes på en population som kan anses löpa hög risk för ligamentskador. I två av studierna [6, 7] studerades fotbollsspelare, två studier [3, 5] omfattade basketbollspelare, en studie inkluderade fallskärmshoppare [4]. Fyra av de inkluderade studierna jämförde en grupp med fotledsortos med en kontrollgrupp. I tre av studierna användes semirigida luftkuddeförsedda ortoser. I en av studierna [7] inkluderades också bruk av ett specifikt träningsprogram med balansplatta. I samtliga studier följdes patienterna upp under en kort period (mindre än ett år).

Metodologisk kvalitet

Den metodologiska kvaliteten varierade. Endast en studie rapporterade hur randomiseringen hade gått till, ingen var dubbelblind eller gav tillräckliga data för att bedöma om randomisering hade skett på adekvat sätt. Medelkvalitetspoäng var 18,2 [15-24] av högsta möjliga 33. Två studier [3, 7] rapporterade patienter som inte kunde följas.

Samtliga studier hade allvarliga »intention to treat«-problem, dvs ingen av studierna beskrev behandlingsprogrammet på ett tillräckligt tydligt sätt. Utvärdering gjordes av tränare i tre av studierna [3, 6, 7], av en läkare (oklart om oberoende) i de två andra studierna. Uppföljningen var således inte blindad i någon av studierna.

II Resultat

Antalet fotledsstukningar både i interventions- och i kontrollgrupperna visade signifikant minskning både när fotledsortos användes (oddsratio, OR, 0,47; 95 procents konfidensintervall 0,35–0,63) och med balansplatta (OR 0,30; 95 procents konfidensintervall 0,15–0,60). Vid både fotledsortos och balansträning var effekten större hos dem som tidigare hade ådragit sig fotledsstukning.

Specifik analys vad gäller yttre stöd versus kontroller visade en signifikant minskning av antalet fotledsstukningar hos individer som tidigare hade haft stukning (OR 0,31; 95 procents konfidensintervall 0,19–0,49). Hos dem som inte tidigare hade haft fotledsstukning var minskningen dock inte signifikant.

Det finns alltså bevis för att fotledsstöd (ortos) har skadeförebyggande effekt hos individer med tidigare skada. Det förelåg ingen skillnad i skadornas allvarlighetsgrad, dvs grad I, II eller III. Ej heller fanns någon skillnad vad gäller risken för övriga skador, speciellt knäskador.

Slutligen gjordes ett försök att jämföra en intervention med en annan. Någon jämförelse av t ex användning av semirigida ortoser versus balansträning var dock inte möjlig på grund av metodologiska problem.

II Diskussion

Av flera studier visade sig endast fem möjliga att inkludera som grund för preventiva åtgärder med hänsyn till vetenskaplig kvalitet. Med undantag för balansbräda, som testades

av Tropp och medarbetare [7], visade samtliga förebyggande interventioner positiv effekt av yttre fotledsstöd. Detta var speciellt markant vad gäller luftkuddeförsedda stöd. Detta stöd består av ett yttre plasthölje som hålls ihop med ett hålband och ett kardborreband runt fotleden (Figur 1). På insidan av plasthöljet finns två luftkuddar, som ökar stödets komfort.

Studierna bedömdes alla vara av endast modest metodologisk kvalitet. Ingen av dem presenterade någon kostnadsanalys, ej heller någon beskrivning av hur behandlingen accepterades, dess effekt på idrottsutövning eller konsekvenserna av användning av fotledsortos på längre sikt. Samtliga dessa faktorer är viktiga om fotledsortos på vetenskapliga grunder skall kunna rekommenderas för idrottsutövare. Särskilt måste kostnaderna beaktas.

Ortos ger god preventiv effekt

Resultatbeskrivningen, särskilt i relation till övriga skador, var svag, men trots dessa begränsningar var samtliga studier väl genomförda med noggrann monitorering av resultatvariabler inom de begränsningar som nämnts ovan. Resultaten pekar starkt på att användning av semirigida ortoser ger en statistiskt signifikant reduktion av antalet fotledsstukningar hos aktiva utövare av högriskidrotter. Subgruppsanalys visade att den preventiva effekten var störst hos individer som tidigare hade ådragit sig ledbandsskador. Det återstår att se om en liknande effekt kan uppnås hos individer som inte tidigare har ådragit sig ledbandsskador i fotleden.

På basis av dessa studier är det oklart huruvida föreliggande åtgärder skulle kunna vara relevanta för andra åldersgrupper, vid andra typer av aktiviteter eller för idrottande kvinnor. Att ge råd om vilka specifika fotledsortoser som bör användas är inte heller möjligt. Å andra sidan är resultaten homogena, vilket pekar på att det är bättre att använda någon form av yttre stöd än inget alls.

Balansträning särskilt anpassad för individer som tidigare ådragit sig ledbandsskador har också bevisad effekt [7]. Ytterligare studier är dock nödvändiga innan allmängiltiga konklusioner kan dras angående förebyggande effekt av sådan träning. Det föreligger också otillräcklig eller ingen vetenskaplig information vad gäller förebyggande effekt av styrketräning.

Konklusion

Det finns ett bra vetenskapligt stöd för användning av yttre fotledsstöd, såsom semirigid ortos eller »air-cast«-ortos, i syfte att förhindra ledbandsskador i fotleden, särskilt laterala skador, i samband med högriskaktiviteter som fotboll, basketboll och volleyboll. Individer som tidigare har ådragit sig en fotledsstukning bör informeras om att risken för nya stukningar kan minskas med användning av yttre fotledsstöd.

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Kannus P, Renström P. Treatment for acute tears of the lateral ligaments of the ankle. *J Bone Joint Surg Am* 1991;73(2):305-12.
2. Quinn K, Parker P, de Bie R, Rowe B, Handoll H. Interventions for preventing ankle ligament injuries (Cochrane Review). In: *The Cochrane Library*. Issue 2, 2001. Oxford: Update Software.
3. Barrett JR, Tanju JL, Drake C, Fuller D, Kawasaki RI, Fenton RM. High- versus low-top shoes for the prevention of ankle sprains in basketball players. A prospective randomized study. *Am J Sports Med* 1993; 21(4):582-5.
4. Ryan JB, Amoroso PJ, Jones BH, Bickley BT, Taylor DC, Leitschuh

- P. Impact of an outside-the-boot ankle brace on sprains associated with military airborne training. *Orthopaedic Transactions* 1994; 18(2):557.
5. Sitler M, Ryan J, Wheeler B, McBride J, Arciero R, Anderson J, et al. The efficacy of a semi-rigid ankle stabilizer to reduce acute ankle injuries in basketball. A randomized clinical study at West Post. *Am J Sports Med* 1994;22:454-61.
 6. Surve I, Schweltnus MP, Noakes T, Lombard C. A fivefold reduction in the incidence of recurrent ankle sprains in soccer players using the Sport-Stirrup orthosis. *Am J Sports Med* 1994;22(5):601-6.
 7. Tropp H, Askling C, Gillquist J. Prevention of ankle sprains. *Am J Sports Med* 1985;13(4):259-62.
 8. Gross MT, Bradshaw MK, Ventry LC. Comparison of support provided by ankle taping and semirigid orthosis. *J Orthop Sports Phys Ther* 1987;9:33-9.
 9. Hughes LY, Stetts DM. A comparison of ankle taping and a semi-rigid support. *Physician and Sportsmedicine* 1983;11:99-103.
 10. Bennett KL, Goldie PA. The differential effects of external ankle support on postural control. *J Orthop Sports Phys Ther* 1994;20(6):287-95.
 11. Burks RT, Bean BG, Marcus R, Barker HB. Analysis of athletic performance with prophylactic ankle devices. *Am J Sports Med* 1991;19(2):104-6.
 12. Robinson JR, Frederick EC, Cooper LB. Systematic ankle stabilization and the effect on performance. *Medicine and Science in Sports and Exercise* 1986;18(6):625-8.
 13. Garrick J, Requa R. Role of external support in the prevention of ankle sprains. *Medicine and Science in Sports and Exercise* 1973; 5(2):200-3.

SUMMARY

Ankle braces prevent ligament injuries

Jon Karlsson

Läkartidningen 2002;99:3486-9

The Cochrane collaboration has performed a meta-analysis of all studies found on the prevention of ankle ligament injuries, frequent in sports like soccer, European handball and basketball. Interventions include the use of modified footwear and associated supports, training programmes and health education. Five randomized trials totalling 3 954 participants were included. With the exception of ankle disc training, all prophylactic interventions entailed the application of an external ankle support in the form of a semi-rigid orthosis, air-cast or high top shoes. The studies showed a significant reduction in the number of ankle sprains in individuals allocated to external ankle support. This reduction was greater for those with a previous history of ankle sprains.

Correspondence: Jon Karlsson, Dept of Orthopaedic Surgery, Sahlgrenska Universitetssjukhuset, SE-416 85 Göteborg, Sweden (jon.karlsson@vgregion.se)

Särtryck

Läkartidningen

När konsensus saknas om hur läkaren bör behandla, spelar den beprövade erfarenheten stor roll. Det 48-sidiga häftet innehåller 32 korta, praktiskt inriktade artiklar med anknytning till vårdens vardag och vänder sig till alla kliniskt verksamma läkare. Förutom diagnostik med terapi speglas goda exempel på prevention, ledningsfrågor och administration.

Priset är 45 kronor

Enligt min erfarenhet



Beställer härmed.....ex
av "Enligt min erfarenhet"

.....
namn

.....
adress

.....
postnummer

.....
postadress

**Insändes till Läkartidningen
Box 5603
114 86 Stockholm**

Faxnummer: 08-20 74 35

www.lakartidningen.se
under särtryck, böcker



**EVIDENSBASERADE STUDIER
I KORTHET**

ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Vitamin B-6 förbättrar sannolikt premenstruellt syndrom och depression

Premenstruellt syndrom är ett svårvärderat tillstånd med en stor mängd behandlingar i omlopp. En översikt från 1980-talet identifierade 327 olika behandlingar. SSRI-preparat används numera en del i Sverige medan inställningen till pyridoxin (vitamin B-6) är

krav som kan ställas på en systematisk översikt. Av 25 publicerade studier var endast 9 randomiserade, dubbelblinda och placebokontrollerade och flertalet uppvisade metodologiska svagheter. Sammantaget inkluderades 940 patienter. Pyridoxin 100 mg per dag förbättrade de premenstruella symtomen och möjligen fanns en effekt också på 50 mg. I fyra studier med 541 patienter påvisades också en förbättring av depressiva symtom.

Författarna är försiktiga i sina slutsatser med tanke på den låga vetenskapliga kvaliteten men hävdar att pyridoxin i doser upp till 100 mg per dag sannolikt minskar premenstruella symtom och depression utan någon påvisbar toxicitet. Noteras bör dock att doser på mer än 200 mg/dag är förenliga med perifer neuropati och nyligen begränsade de brittiska hälsovårdsmyndigheterna doserna som säljs via apotek till högst 50 mg/dag.

Långtidsstudier med tillräckligt antal patienter för att belysa såväl nytta som skada med behandlingen är nödvändiga liksom jämförelser med antidepressiva av SSRI-typ.

Wyatt KM, Dimmock PW, Jones PW, Shaughn O'Brien PM. Efficacy of vitamin B-6 in the treatment of premenstrual syndrome: systematic review [see comments]. *BMJ* 1999; 318(7195): 1375-81.

avvaktande trots att preparatet används en del via hälsokostaffären.

För att klarlägga frågan om pyridoxin kan förbättra symtomen vid premenstruellt syndrom genomförde engelska gynekologer en översikt och metaanalys. Arbetet är utfört med stor noggrannhet och med iakttagande av alla de

Ännu inga belägg för att av homocystein leder till

Systematiska översikter med åtföljande metaanalyser (där så är lämpligt) är kraftfulla verktyg för att sammanställa medicinsk kunskap med minsta möjliga systematiska fel (bias). De flesta sådana översikter berör nyttan och riskerna med olika typer av interventioner, värderade i randomiserade och kontrollerade studier. Endast enstaka systematiska översikter över diagnostiska metoder och epidemiologiska frågeställningar har publicerats.

Frågan om kausalsamband mellan en faktor och utvecklande av sjukdom är central både inom kardiovaskulär och cancerforskning och är sällan möjlig att definitivt klarlägga utan interventionsstudier mot den misstänkta riskfaktorn. Vi har kunnat följa debatten om hyperinsulinemi, typ A-personlighet och senaste åren även förhöjda nivåer av homocystein i blodet som markörer för, parallella fenomen, epifenomen eller kausala orsaker till kardiovaskulär sjukdom.

I en rykande färsk systematisk översikt har alla de 43 epidemiologiska studier som studerat sambandet mellan homocysteinemi och hjärtinfarkt eller stroke inkluderats. De flesta fall-kon-

Effektiv smärtlindring och behandling mot illamående vid dagkirurgi – en medicinsk utvärdering

För att framgångsrikt bedriva dagkirurgi bör man kunna erbjuda en effektiv postoperativ smärtbehandling och minimera illamående och kräkning.

Smärtforskarna och anesthesiologerna Henry McQuay och Andrew Moore från Radcliffe Hospital i Oxford (centralfigurer i Centre for Evidence-Based Medicine och redaktörer för den internetbaserade tidskriften i samma ämne, *Bandolier*), genomförde nyligen en medicinsk utvärdering för att ge underlag om bästa val av analgetika och antiemetika postoperativt.

Klart ineffektiva som smärtlindrande klassades transkutan nervstimulering, lokala morfinitinjektioner och kodein i doser om 30 mg. Att spruta morfin i en opererad knäled ger ingen kliniskt relevant analgesi och att ge NSAID-preparat i injektioner eller rek-

talt ger ingen bättre effekt (men innebär möjligen större risker) än peroral behandling med samma preparat.

Givna som standarddoser har dextropropoxifen, tramadol, paracetamol, ibuprofen och diklofenak ett bevisat värde. De två sistnämnda har ekvivalent smärtstillande effekt som 10 mg morfin intramuskulärt. Kombinationer paracetamol och kodein är höggradigt effektivt. Trots att topiskt applicerat NSAID fungerar bra mot smärre skador och kronisk smärta fyller de ingen funktion vid dagkirurgi.

Illamående och kräkning postoperativt kan antingen hanteras genom förebyggande behandling till samtliga eller behandling enbart av dem som uppvisar tecken på illamående. Inget av de preparat som studerades var tillräckligt

effektivt för att användas i förebyggande syfte. En kostnadseffektanalys visade att profylax med nyare antiemetika inte förhindrade mer illamående och kräkning än att behandla dem som redan blivit illamående eller kräkt. En sådan profylaktisk behandling utsatte däremot patienterna för betydligt mer läkemedel till avsevärt högre kostnad.

Smärtforskarna pekar på behovet av att utforska andra komponenter i hela det dagkirurgiska omhändertagandet än narkosmedel och antiemetika för att minska problemet med illamående.

McQuay HJ, Moore RA. Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review. *Health Technol Assess* 1998; 2(12): 1-236.

förhöjda blodnivåer kardiovaskulär sjukdom

trollstudier liksom tvärsnittsstudier finner förhöjda blodnivåer av homocystein hos personer med känd kärlsjukdom jämfört med personer utan sjukdom.

Sambandet kvarstår även efter det att man tagit hänsyn till ett antal potentiella störfaktorer (confounders). I de sju prospektiva studierna avseende hjärtinfarkt finner man inget samband mellan homocysteinnivåer vid studiens start och utvecklande av sjukdomen under många års uppföljning. Undantaget är två studier där patienter med existerande kärlsjukdom inkluderades.

Författarnas slutsats är att förhöjda nivåer av homocystein kan vara en akutfas-reaktant (som tidigare spekulerats kring såväl fibrinogennivåer och fibrinolytiska faktorer) som skall ses som en markör för existerande ateroskleros. Författarna hävdar att vi måste invänta de randomiserade och kontrollerade studierna där homocysteinnivåer sänks och kardiovaskulär sjukdom analyseras innan frågan om homocysteinemi som kausal riskfaktor kan besvaras.

Christen WG, Ajani UA, Glynn RJ, Hennekens CH. Blood levels of homocysteine and increased risks of cardiovascular disease: causal or casual? Arch Intern Med 2000; 160(4): 422-34.

Nasala steroider bättre än antihistaminer vid allergisk rinit

Förekomsten av allergisk rinit har fördubblats i Sverige de senaste trettio åren. Sjukdomen är inte bara förenad med lokala obehag utan kan allvarligt nedsätta välmående och livskvalitet. Egen medicinerings med antihistaminer är omfattande men det har funnits motstridiga uppfattningar om nasala steroider är bättre än antihistaminer.

I en mycket välgjord metaanalys inkluderades 2 267 personer från 16 randomiserade,

dubbelblinda studier och resultaten talade starkt till förmån för nasala steroider. Dessa var mer effektiva än peroral antihistaminer avseende symptom som nästäppa, rinnsnuva, nysningar och även vid en global bedömning av försämringsgrad men man fann ingen skillnad för ögonsymtom.

Att ange skillnader i absoluta tal är svårt då många olika mått använts men ett typiskt fynd är att patienter som be-

FOTO: PRESSENS BILD



handlats med steroider nasalt fick 1,8 extra symptomfria dagar per vecka jämfört med de patienter som fått antihistaminer. Inga studier inkluderade livskvalitetsmått som är väsentliga med tanke på den trötthet och nedsättning av allmäntillståndet som många patienter med allergisk rinit upplever. Utifrån en enkel analys, baserad på australiska förhållanden, ansåg författarna att nasala steroider var mer kostnadseffek-

tiva. Man anger att de nasala steroiderna bör vara förstahandspreparat men att antihistaminer kan ha en roll, speciellt vid irritation från ögonen som inte viker för steroider.

Weiner JM, Abramson MJ, Puy RM. Intranasal corticosteroids versus oral H1 receptor antagonists in allergic rhinitis: systematic review of randomised controlled trials. BMJ 1998; 317(7173): 1624-9.

Interaktivt lärande förändrar läkares beteende

Läkares fortbildning förväntas leda till förbättrad klinisk förmåga och därmed till bästa tänkbara hälsoeffekter för deras patienter. En noggrann översikt över det medicinska fortbildningssystemet (continuing medical education, CME) som inkluderade 14 nordamerikanska, randomiserade studier fann ingen effekt av traditionella föreläsningar vare sig mätt som förändringar i klinisk praxis eller som patientnytta.

Med ett interaktivt lärande såsom rollspel, diskussionsgrupper, praktisk träning, problemlösning eller fallstudier kunde förändringar i läkarnas beteende påvisas och i enstaka fall också förbättringar i patientrelaterade mått. Att dela upp det interaktiva lärandet i flera seanser tycktes också öka genomslagskraften.

Författarna drar slutsatsen att enbart föreläsningar med passiva åhörare inte

FOTO: VEIJO MEHTONEN



leder till önskvärda förändringar i läkares kliniska agerande. Detta utesluter inte att föreläsningar kan ha effekt på andra aspekter på professionell kompetens såsom kunskaper, färdigheter eller attityder och därmed fungera som ett stöd för förändring.

Om vår uppfattning är att fortbildning bör leda till konkret förändring i patientarbetet med åtföljande nytta för

patientens hälsa så bör sådana utfall dock betraktas som en form av surrogat-effektmått.

Davis D, O'Brien MA, Freemantle N, Wolf FM, Mazmanian P, Taylor-Vaisey A. Impact of formal continuing medical education: do conferences, workshops, rounds, and other traditional continuing education activities change physician behavior or health care outcomes? JAMA 1999; 282(9): 867-74.

Kan man förebygga bakteriuri vid KAD?

Många patienter som läggs in på sjukhus får en kvarliggande kateter (KAD) under kortare eller längre tid. Några procent av patienterna utvecklar varje dag bakteriuri, och ca 30 procent får också kliniska symtom, ibland allvarliga infektioner med bakteriemi och sepsis som följd. En systematisk översikt i de amerikanska invärtesmedicinernas tidskrift *Annals of Internal Medicine* tar sig an problemet i hela dess bredd, alltifrån riskfaktorer till katetertyper och farmaka.

Viktigaste faktorn för att utveckla bakteriuri är durationen av kvarliggande kateter, men dessutom är äldre patienter, kvinnor, diabetiker och de med nedsatt njurfunktion mera drabbade. Risken att från bakteriuri utveckla bakteriemi och sepsis tycks vara störst för äldre män med sjukdomar i urinvägarna. Växt av *Serratia* ökade risken betydligt.

Vid långtidsbruk visar flera randomiserade studier att suprapubiska katetrar är förenade med lägre risk för infektioner (och dessutom är kateterbärarna mera nöjda!). Frånsett den vanliga

aseptiska tekniken vid inläggning och skötsel av katetern finns det få åtgärder som bevisat minskar bakterieväxt och infektion. Blåssköljning, antibakteriella medel i urinpåsen eller rigorös rengö-

metaanalys. Systemiska antibiotika är effektiva under korta behandlingstider, men problem med resistens förhindrar sådan behandling.

Metenaminhippurat (i Sverige godkänt för underhållsterapi vid asymtomatisk bakteriuri och KAD) saknar väsentligen modern klinisk dokumentation. Små och icke-randomiserade studier har visat vissa gynnsamma effekter men måste tillmätas tveksamt bevisvärde. Då preparatet dessutom kräver pH under 6 i urinen bör det egentligen tillföras tillsammans med surgörande substanser, vilket sällan sker. Författarna efterlyser stora och randomiserade studier för att slutligen klarlägga om preparatet verkligen är till någon nytta för kroniska kateterbärare.

Med tanke på hur lite det finns att göra för att för-

hindra nosokomiala infektioner vid längre tids KAD-behandling måste slutsatsen bli att använda sådan så lite och så kort tid som möjligt.

FOTO: IBL/BILBYRA



Växt av bakterien *Serratia marcescens* ökar risken för utveckling av bakteriemi och sepsis vid långvarig användning av kateter.

ring med antibiotiska salvor kring meatus har alla misslyckats med att minska infektionerna.

Silveroxidbelagda katetrar skiljer sig inte från vanliga katetrar men beläggning av silverföreningar på katetern leder till avsevärt bättre resultat i en

Arch Intern Med 1999; 159: 800-8.

Systolisk hypertoni – en oberoende riskfaktor för hjärt-kärlsjukdom

Den isolerade systoliska hypertoni är starkt åldersberoende och förekommer hos 25 procent av dem som uppnått 80 års ålder. Det har länge varit omdebatterat huruvida systolisk tryckstegring verkligen medför ökad risk för hjärtinfarkt, stroke eller kardiovaskulär död. Nyttan av behandling har också varit ifrågasatt.

I en systematisk översikt inkluderades åtta studier och drygt 15 000 patienter med medelålder på cirka 70 år. En av de åtta studierna (Syst-China) var inte randomiserad utan fördelade patienterna alternerande till placebo eller aktiv behandling.

Denna studie inkluderas bara i ana-

lysen av det systoliska blodtryckets roll som riskfaktor.

I en så kallad epidemiologisk analys finner man att för varje ökning av det systoliska blodtrycket med 10 mm Hg ökade insjuknandet i stroke med 22 procent och den totala dödligheten med 26 procent. Riskökningen var oberoende av ålder, kön och diastoliskt blodtryck. Däremot noterades ingen ökning av koronara händelser.

De sju behandlingsstudierna använde tiazider, betablockerare och kalciumblockerare som förstahandspreparat. Total mortalitet minskade inte signifikant med blodtrycksbehandling under fyra år, men den kardiovaskulära

dödligheten minskade med 16 procent (1–28 procent). Insjuknandet i stroke minskade med 30 procent (17–41 procent) och infarkter (dödliga och icke-dödliga) med 23 procent (10–35 procent).

Nyttan av behandling i absoluta tal var större för män över 70 år och med känd hjärtkärlsjukdom. Författarna hävdar att aktiv blodtryckssänkning är berättigad hos äldre människor med systoliska blodtryck över 160 mm Hg, i första hand för att reducera insjuknande i stroke.

Lancet 2000; 355: 865-72.

Svensk frontlinjeforskning

Inga gynnsamma effekter av långtidsbehandling med hepariner vid akuta koronara syndrom

Eikelboom JW, Anand SS, Malmberg K, Weitz JI, Ginsberg JS, Yusuf S. Unfractionated heparin and low-molecular-weight heparin in acute coronary syndrome without ST-elevation; a meta-analysis. *Lancet* 2000; 355:1936-42. (Kommentar: Becker RC. Heparins in management of acute coronary syndromes without ST-segment elevation. *Lancet* 2000; 355: 1926-7.)

Lågmolekylära hepariner, som tillägg till ASA, har fått en stor användning vid akuta koronara syndrom utan ST-höjning (instabil angina och icke-Q-våg infarkter). För att klarlägga effekten av behandlingstidens längd och jämföra de lågmolekylära heparinerna med standard heparinbehandling utförde vi en systematisk översikt med metaanalys. En extensiv sökning efter publicerade och opublicerade studier identifierade 12 randomiserade och kontrollerade studier med drygt 17 000 patienter.

Under de första sju dyggen minskar heparin eller lågmolekylärt heparin, i tillägg till ASA, risken för död eller infarkt från 7,4 procent till 4,5 procent (odds-ratio 0,53, 95 procents konfidensintervall 0,38–0,73). Det innebär att att behandlingen förebygger 29 sådana händelser, huvudsakligen icke-fatala infarkter, om 1 000 patienter behandlas under en vecka.

I fem studier reduceras död eller infarkt från 2,3 till 2,2 procent med lågmolekylärt heparin jämfört med vanligt heparin (odds-ratio 0,88, 0,69–1,12). Sammanlagt 11 000 patienter ingår i en jämförelse av långtidsbehandling med lågmolekylärt heparin mot placebo.

Vi påvisade ingen minskning av död eller hjärtinfarkt efter sju dagars behandling men däremot en signifikant ökad risk för allvarliga blödning-

ar (odds-ratio 2,26, 1,63–3,14). Om vi behandlar 1 000 patienter under längre tid än en vecka kommer det att uppstå 12 allvarliga blödningar.

Således är det klart att kort tidsbehandling med hepariner (oavsett typ) halverar risk för död eller infarkt vid instabil kranskärlsjukdom men det finns inga övertygande belegg för någon skillnad mellan heparin och lågmolekylärt heparin. Dock är ju lågmolekylärt heparin betydligt mer lättanvänt vilket sannolikt också ökar tillgängligheten. Vidare finns till dags dato inga övertygande bevis för att långtidsbehandling med lågmolekylärt heparin gör ytterligare nytta.

Klas Malmberg

Kardiologkliniken,

Karolinska sjukhuset samt

Mc Master University, Hamilton, Kanada

Regional eller generell anestesi vid höftfraktur?

Varje år opereras många tusen äldre patienter med höftfraktur, numera nästan enbart i regional anestesi. Med spinal/epidural anestesi är den perioperativa handläggningen enklare jämfört med generell anestesi, som upplevs som en större belastning för den äldre och oftast multipelt sjuka patienten.

De senaste åren har man dock uppmärksammat problemen med regional anestesi vid t ex instabil kranskärlsjukdom då det snabba blodtrycksfallet kan vara skadligt. Riskerna för blödning i spinal-/epiduralrummet i samband med tromboprofylaktisk antikoagulation har nyligen aktualiserats från Läkemedelsverket. Det har hitintills saknats en övertygande evidens om vilken narkosform som är den bästa när patientrelaterade variabler vägs in i beslutet.

Brittiska forskare inom Cochrane-samarbetet presenterade nyligen en systematisk översikt där effekterna av de två narkosformerna jämförs. Femton studier, bland annat japanska studier på originalspråket, med 2 162 patienter granskades. Studierna var, med två un-

dantag, små och behäftade med metodologiska problem.

Dödligheten efter en månad var 6,4 procent i gruppen med regional anestesi jämfört med 9,4 procent i gruppen med generell anestesi (odds-ratio 0,66, 95 procents konfidensintervall 0,46–0,96). Efter tre månader var skillnaden i dödlighet inte längre signifikant. Endast två studier rapporterar data efter ett år då cirka 22 procent av patienterna avlidit i båda grupperna.

Författarna konkluderar att regional anestesi medför marginella fördelar, i första hand en något lägre mortalitet efter en månad. De spekulerar över att regional anestesi möjliggör att den medicinskt sköra patienten överlever ingreppet för att inom kort tid därefter avlida till följd av sina andra medicinska problem.

Parker MJ, Urwin SC, Handoll HHG, Griffiths R. General versus spinal/epidural anaesthesia for surgery for hip fractures in adults (Cochrane Review). In: *The Cochrane Library*, Issue 2, 2000. *Br J Anaesth* 2000; 84: 450-5.

Skriv till oss!

Läkartidningen vill uppmantra läsarna att själva skriva kommentarer och referat av systematiska översikter med relevans för en bredare läsekrets. Var fjärde vecka kommer vi att avsätta upplaget »Nya rön» för evidensbaserad medicin med betoning på referat av bra systematiska översikter och metaanalyser. Andra aktuella artiklar som berör evidensbaserad medicin kan också vara av intresse för tidningens läsare. Om du har läst något som du vill bibringa läsekretsens, kontakta Mats Eliasson, medicinsk redaktör för evidensbaserad medicin:

Mats.Eliasson@nll.se
eller 070-513 02 93.

Systematiskt utvecklade norska riktlinjer för UVI hos kvinnor

Akuta urinvägsbekymmer hos kvinnor är en vanlig besöksorsak i primärvården och nyligen publicerade norska riktlinjer är därför välkomna. Riktlinjerna, som utgår från Statens institutt for folkehelse, gäller friska icke-gravida kvinnor i åldern 16–55 år med akuta, distala urinvägsbesvär. De bygger på en systematisk litteratursökning i Medline, Cochranebiblioteket och andra källor och inkluderade enbart dokumentation som uppfyllde uttalade kvalitetskrav. Riktlinjerna genomgick därefter en föredömlig remissbehandling av laboratoriemedicinare, allmänmedicinare och andra specialister, allmänmedicinska institutioner och vetenskapliga föreningar samt av mottagningspersonal och patienter.

Sannolikheten för bakterieuri är omkring 80 procent för en kvinna med ty-

piska symtom på blåskatarr. Om såväl nitrit- som leukocyttest på sticka är negativa kommer sannolikheten för urinvägsinfektion, UVI, ändå att vara omkring 50 procent. Antibiotikabehandling tycks ge snabbare symtomlindring, men man fann begränsat stöd för detta från placebokontrollerade studier. Många kvinnor söker inte läkare vid symtom på UVI.

Riktlinjerna anger att kvinnor med typiska symtom kan få antibiotika utan undersökning av urinen. Kvinnor som tidigare haft UVI kan erbjudas råd och behandling per telefon. Däremot bör de kvinnor som inte tidigare haft UVI och de med atypiska besvär erbjudas läkarbesök. Vid okomplicerad UVI är antibiotikabehandling i tre dagar tillräckligt, baserat på bakterieeliminering. Distal UVI är normalt ett självbegrän-

sande tillstånd, varför den som så önskar kan avvakta utan antibiotikabehandling. Den norska bedömningen är att implementering av riktlinjerna kan leda till bättre service åt den stora gruppen kvinnor med UVI.

Slutsatserna bör äga giltighet även för svenska förhållanden, men skulle möjligen kräva en noggrann uppföljning av antibiotikaföreskrivning på denna indikation, ytterligare uppstramning av aktuell kinolonföreskrivning samt övervakning av resistensutvecklingen hos urinvägspatogener bland samhällsförvärvad UVI.

Anders.Hernborg
Anders.Hernborg@telia.com

Tidskr Nor Laegeforen 2000; 120: 1748-53.
<http://www.legeforeningen.no/tsweb/> (sök i arkivet på författarnamnet Flottorp)

Kostnadseffektivitet vid behandling och sekundär prevention av stroke

Terapier med stora behandlingseffekter (både i absoluta och relativa tal) kan ha en blygsam effekt på sjukdomens utfall i hela populationen om bara en liten del av alla sjuka kan komma in fråga för interventionen. Ett typiskt exempel på detta är trombolysbehandling vid akut ischemiskt stroke som kan ges till som mest 5–10 procent av strokepatienter. Å andra sidan kan interventioner med små effekter få ett stort genomslag vad gäller invaliditet och död, om flertalet patienter kan åtnjuta behandlingen såsom är fallet med acetylsalicylsyra (ASA) vid sekundärprevention av ischemiskt stroke eller TIA. Slutligen kan vi i en trängd sjukvårdsekonomi inte blunda för kostnadsjämförelser mellan olika behandlingalternativ vid ett givet tillstånd.

En föredömligt pedagogisk och lättförståelig systematisk översikt visar G Hankey och C Warlow vilka interventioner som har ett rimligt vetenskapligt stöd vid akut stroke och sekundärprevention därefter. Samtliga kalkyler redovisas utifrån en befolkning på en miljon där man kan förvänta sig 1 800 förstagångsinsjuknanden och 600 recidiv per år. Man kan under ett års uppföljning förvänta sig 700 dödsfall och 600 som överlever med hjälpbehov. Dessa 1 300 patienter är de som man i teorin skulle kunna hjälpa med effektiva interventioner. Kostnader beräkna-

des från australiska förhållanden.

Vid det akuta insjuknandet reducerar både behandling vid strokeenhet och trombolys död och resttillstånd i lika stor utsträckning (absolut riskreduktion 6 procent, NNT 17). Men eftersom strokeenheter kan göras tillgängliga för uppåt 80 procent av patienterna, och trombolys som mest kan ges till 10 procent blir antalet patienter som räddas 107 respektive 15! Strokeenheten ansågs inte medföra extrakostnader, medan kostnaden per patient som räddas från död eller invaliditet med trombolys med tPA blir 180 000 kr. ASA i dos 300 mg kan ges till nästan samtliga patienter med akut ischemiskt stroke och trots en blygsam reduktion av absolut risk på 1,2 procentenheter räddas ändå hela 23 patienter till en kostnad av 400 kr per patient. Heparin i akutfasen finns inget stöd för.

Nyttan av sekundärprevention beräknas på den grupp på 12 000 i befolkningen som överlevt ett stroke eller TIA och som har en årlig risk för recidiv på 7 procent. De absoluta riskreduktionerna är genomgående i storleksordningen 1–2,5 procentenheter för blodtrycksbehandling, rökstopp, statinbehandling av hyperkolesterolemi och olika trombocythämmare (ensamma eller i kombination). Ett klart undantag är antikoagulantia vid förmaksflimmer som reducerar risken från 12 till 4 procent. Vissa in-

terventioner blir mycket kostnadseffektiva då de kan erbjudas en stor andel patienter till lågt pris såsom tiazidbehandling av hypertoni, ASA, kort rådgivning för rökstopp och antikoagulantia vid förmaksflimmer. 10–20 gånger högre blir kostnaden per patient som räddas från recidiv med statiner eller ASA + dipyridamol. Behandling med klopidogrel ger femtio gånger högre kostnader och slutligen ger karotiskirurgi cirka 100 gånger högre kostnad per uteblivet recidiv i befolkningen. Författarna drar slutsatsen att strokeenheter är högst prioriterade tillsammans med ASA samt antikoagulation vid flimmer, blodtrycksbehandling och rökstopp.

I ett svenskt perspektiv kan vi notera att strokeenheter trots allt har något högre bemanning och kostnader än vanliga avdelningar, men kostnadsjämförelser torde fortfarande utfalla mycket gynnsamt. För att öka kostnadseffektiviteten av de dyrare interventionerna måste vi få kunskap om vilka som drar mest nytta av dem. Vi bör också bejaka den vånda beslut om trombolys kan innebära med en omedelbart ökad risk för blödningskomplikationer hos några patienter för att rädda andra patienter från livslånga handikapp.

Lancet 1999; 354: 1457-63.
<http://www.thelancet.com/newlancet/sub/issues/vol354no9188/review1457.html>

Dåligt visad nytta av behandling i tryck-kammare vid kolmonoxidförgiftning

Standardbehandling vid förgiftning med kolmonoxid är tillförsel av 100 procent syrgas utan övertryck (normobar oxygenbehandling) för att öka elimineringen av kolmonoxid ur kroppen. Genom att tillföra syrgas med högre tryck i en tryckkammare (hyperbar oxygenbehandling, HBO) kan elimineringen påskyndas ytterligare. Avsikten med behandlingen är huvudsakligen att förhindra persisterande respektive sent insättande neurologiska skador. Eftersom tryckkammare bara är tillgänglig vid sex sjukhus i landet kan inte alla som exponerats för kolmonoxid behandlas med HBO. Svenska riktlinjer anger att tryckkammarbehandling är indicerad då patienten har varit fortfarande är medvetslös. Även kvarstående neurologiska symtom under normobar behandling anses motivera HBO.

Inom Cochranesamarbetets grupp för trauma och skador har kanadensiska forskare granskat underlaget för HBO vid kolmonoxidförgiftning. Totalt fann man sex randomiserade studier, men bara tre kunde ligga till grund för en analys där förekomst av neurologiska symtom efter en månad valdes som det viktigaste patientrelaterade effektmåttet. Kontrollgrupperna behandlades med normobar syrgasbehandling. Studierna inkluderade 455 patienter med varierande förgiftningsgrad och olika sätt att genomföra såväl den hyperbara som normobara oxygenbehandlingen.

En metaanalys av de tre studierna visade att förekomst av neurologiska symtom i gruppen som behandlats i tryckkammare var 34,2 procent mot 37,2 i gruppen som fått konventionell normobar oxygenbehandling. Den absoluta riskreduktionen på 3 procenten-

heter var inte statistiskt signifikant och kan också uttryckas som ett odds-ratio på 0,82 (95-procentigt konfidensintervall 0,41–1,66), dvs resultatet är inte heller helt oförenligt med en viss ökad risk för neurologiska skador med HBO. Intressantast för en svensk läsare är nog resultatet av den enda studien som stratifierats utifrån förekomst av medvetslöshet eller ej. I denna studie som inkluderade flera hundra patienter och analyserades enligt »intention to treat» fann man ingen skillnad i förekomst av neurologiska symtom efter hyperbar behandling, oavsett om patienterna varit medvetslösa eller ej.

Cochranerapporten, som uppdaterades så sent som februari 2000, ger inget stöd för ett allmänt bruk av HBO vid kolmonoxidförgiftning, men viktigare är att det inte heller tycks finnas något gediget vetenskapligt stöd för sådan behandling vid medvetslöshet – den i Sverige dominerande indikationen. Det verkar vara dags för en randomiserad multicenterstudie i Sverige där enbart medvetslösa patienter randomiseras till normobar eller hyperbar oxygenbehandling. Självklart skall uppföljning ske över lång tid med avseende inte bara på neurologiska och kognitiva funktioner utan dessutom på hälsorelaterad livskvalitet och andra mått med relevans för individen som drabbats. Med tanke på det bristande underlaget kan etiska kommittéer knappast lägga hinder i vägen.

Juurlink DN, Stanbrook MB, McGuigan MA. Hyperbaric oxygen for carbon monoxide poisoning (Cochrane Review). In: The Cochrane Library, Issue 2, 2000. Oxford: Update Software (nås via www.slf.se/DoktorsSida)

Från evidensbaserad kunskap till beslut om min enskilda patient

Den vanligaste kommentaren jag möter vid föreläsningar om evidensbaserad medicin är: »Visst, det är utmärkt att få en bättre kunskapsbas att stå på och veta vad som fungerar eller inte. Men denna kunskap berör grupper av patienter som är rekryterade till studier på grund av speciella karakteristika. Vad jag nu behöver är ett instrument för att gå från det generella till det specifika, dvs till min enskilda patient. Hur kan jag omvandla graden av nyttan och väga in oönskade effekter av behandlingen för min patients speciella demografi och sjukdomsbild? Hur väger jag in patientens värdering av olika utfall (skulle patienten hellre bli liggande förlamad i emboliskt stroke än avlida i en Waranblödning?) i ett beslutsunderlag?»

Frågorna berör många av den kliniska epidemiologins kärnämnen och omfattar bl a att försöka kalkylera min patients individuella och specifika absoluta risk för att drabbas av det vi vill undvika med behandlingen. Det handlar också om att kvantitativt justera den förväntade behandlingens vinsten till min patients ålder, kön, samtida sjukdom osv. Dessa och många andra frågor diskuteras utifrån ett fiktivt exempel i en rykande färsk artikel i serien Users' Guides to the Medical Literature (se LT 26/2000).

Jag kommer att återvända till de viktiga men svåra frågorna under den fortsatta serien om evidensbaserad medicin i Läkartidningen men vill ge de intresserade läsarna chansen att redan nu kasta sig över denna läsvärda artikel.

JAMA 2000; 283: 2829-36.
<http://jama.ama-assn.org/issues/>

Icke-signifikant är inte detsamma som likvärdigt!

Det har blivit allt vanligare med studier som »bevisar» att en ny behandling är likvärdig med en tidigare använd behandling. Så kan till exempel en ny tryptan, statin eller bifosonat beskrivas som likvärdig med »original»-substan- sen. Sådana ekvivalenskrav baseras oftast på att man i randomiserade, kontrollerade studier inte kunde »avvisa nollhypotesen», dvs vid en jämförelse mellan preparaten var skillnaden i effekt »inte statistiskt signifikant», oftast på 5-procentsnivån ($P > 0,05$). Sådana ekvivalensresonemang möter vi också vid jämförelser mellan olika inhalationshjälpmedel vid astma, jämförelser mellan inhalationssteroider osv. Avsik-

ten är oftast att visa att den nya produkten eller beredning, som kan vara billigare eller lättare att använda därför bör användas – »de är ju ändå lika bra».

Alvar Feinstein, redaktör för Clinical Epidemiology, har nyligen beskrivit vad som krävs för att hävda ekvivalens och studerat 88 artiklar som uppgavs testa ekvivalens mellan två interventioner. I hela två tredjedelar av studierna baserades detta på att nollhypotesen inte kunde avfärdas (dvs $P > 0,05$). Mycket få studier använde sig av modellen att a priori specificera en skillnad i resultat som skulle betraktas som liten och kliniskt oviktig samt därefter testa utfallet mot detta med hjälp av konfi-

densintervall eller liknande. En adekvat kalkyl av en studies »power» krävs också för att kunna diskutera skillnader mellan typ 2-fel och ekvivalens och man finner mycket riktigt att studier som hävdar ekvivalens ofta är gravt underdimensionerade. Artikeln är en viktig väckarklocka för läkemedelskommittéer och andra som allt oftare tvingas ta ställning till jämförelser mellan preparat, men nu inte med påstående om preparatens överlägsenhet utan i stället deras likvärdighet.

Ann Intern Med 2000; 132: 715-22.
<http://www.annals.org/issues/v132n9/full/200005020-00006.html>

Bristande stöd för nytta av patientutbildning vid typ 2-diabetes

Patienter med typ 2-diabetes bör ta aktiv del i behandlingen av sin sjukdom och vårdens uppgift är att ge dem de rätta instrumenten för att klara uppgiften. Denna grundläggande tes har starkt stöd i diabetesvården liksom i patientförbundet. En informerad patient bör kunna spela en aktivare roll i beslut om behandling och därmed också agera rationellt i frågor om kost, motion, rökstopp, viktmedgång eller blodsockertestning. I slutändan skulle sådana beslut om förändrad livsstil kunna ha betydligt större inflytande på överlevnad, komplikationer och livskvalitet än intensiva farmakologiska insatser för förbättrad blodsockerbalans. Trots att erfarenheten har lärt oss att enbart förbättrade kunskaper sällan förändrar beteenden så har många ansett att patientutbildning är en lösning.

Runt om i landet bedrivs diabeteskolor eller diabetesgrupper som har målsättningen att öka kunskap, förändra beteenden, förbättra metabola förhållanden samt stimulera patientansvar och autonomi. Gruppdynamik och interaktivitet samt deltagarstyrd verksamhet betonas. Tyvärr är det mycket sparsamt med bra data avseende effekten av insatserna. De flesta rapporter fokuserar på före-efter-data med kort uppföljningstid, men utan randomiserade kontrollgrupper är det svårt att veta vad som egentligen fungerar och vad som är specifika omhändertagande-effekter och »regression towards the mean».

Det brittiska Centre for Reviews and Dissemination publicerade nyligen en systematisk översikt av insatser för att öka typ 2-patienters förmåga att själva hantera och ta ansvar för sin sjukdom, »promotion of self-management». Totalt identifierades 53 randomiserade och kontrollerade studier, men endast sju redovisade en uppföljning på ett år eller mer. Studierna var små och i analysen inkluderades de 13 undersökningar som omfattade minst 100 patienter. Man ville veta om interventionen förbättrade kunskaper, vikt, HbA1c eller blodtryck. Endast två studier redovisade hälsorelaterad livskvalitet. Generellt var den vetenskapliga kvaliteten låg.

Program som fokuserade på kunskapsförmedling och tränade färdigheter (blodsockertest) visade i flera fall förbättrad långtidkunskap om sjukdomen men inga kliniskt relevanta förbättringar av riskfaktorer. Många av studierna var dessutom genomförda med patienter vars baskunskaper var långt under de svenska patienternas. Program som baseras på inläringsteori redovisas som »kognitiv beteendeförändring» och fokuserar på hälsorelaterat beteende, viktminskning med diet och fysisk aktivitet. Patienten fick sätta upp mål och relatera problemlösning till dessa genom att modifiera sin självuppfattning. I två studier av individuell kognitiv behandling fann man inga förändringar men i en långtidsstudie av gruppbehandling var HbA1c 0,9 procent lägre än i kontrollgruppen och livskvalite-

ten högre. Ytterligare en studie påvisade bestående viktmedgång, men generellt var förändringarna små och heterogena.

Program som vill uppnå »empowerment» av patienterna med typ 2-diabetes uppmuntrade dem att själva söka och välja information, delta aktivt i beslut och »förhandla» med vårdpersonalen. En svensk studie från 1986 visade att tre månaders problemorienterad och deltagarstyrd utbildning gav bättre kunskap efter ett år men ingen skillnad i HbA1c. Inga skillnader i kunskap, hälsobeteende, HbA1c eller vikt kunde påvisas i en aktuell studie av 600 patienter som gavs fritt val av ämnen för utbildning.

Slutsatsen blir att dagens insatser för att stärka patienten med typ 2-diabetes har svagt vetenskapligt stöd. Detta får inte förväxlas med att insatserna inte gör någon nytta men om det verkligen är så kan vi bara spekulera om. Varför inte starta en multicenterstudie med de svenska diabetesenheter som i dag bedriver patientutbildning, randomisera till olika modeller eller sedvanlig behandling? En långtidsuppföljning, som inkluderar inte bara surrogat-effektmått utan även hälsorelaterad livskvalitet och komplikationer, vore nödvändig.

Effective Health Care 2000; 6: 1
www.york.ac.uk/inst/crd/ehc61.htm

Skriv till oss!

Läkartidningen vill uppmuntra läsarna att själva skriva kommentarer och referat av systematiska översikter med relevans för en bredare läsekrets. Vi kommer regelbundet att avsätta uppslaget »Nya rön» för evidensbaserad medicin med betoning på referat av bra systematiska översikter och metaanalyser. Andra aktuella artiklar som berör evidensbaserad medicin kan också vara av intresse för tidningens läsare.

Om du har läst något som du vill bringa läsekretsen, kontakta Mats Eliasson, medicinsk redaktör för evidensbaserad medicin:

Mats.Eliasson@nll.se
eller 070-513 02 93.

COX-2-hämmare summerade

Introduktionen av COX-2-hämmare har väckt stora förhoppningar om effektiv smärtlindring med mindre biverkningar än med NSAID-preparat, men de höga priserna för de nya preparaten celecoxib och rofecoxib stämmer till eftertanke. Landets läkemedelskommitteer har en grannlaga uppgift att formulera behandlingsrekommendationer, men nu rycker den brittiska EBM-tidskriften *Bandolier* ut för att hjälpa till. I en extensiv systematisk översikt granskar man flertalet kliniska aspekter på COX-2-hämning och inkluderar även av tillverkarna tillhandahållet, men ej publicerat, material.

Effekten på akut smärta har jämförts med ibuprofen (400 mg) i tre kontrollerade studier för rofecoxib (50 mg) och en för celecoxib (200 mg). Antalet pa-

tienter som behövde behandlas för att en patient skulle uppnå minst 50 procent smärtlindring, »number-needed-to-treat», NNT, var 1,9 för rofecoxib, 2,8 för celecoxib och 2,1 för ibuprofen.

Motsvarande siffror för paracetamol/kodein (1 000/60 mg) är 1,9; paracetamol (1 000 mg) 4,6; acetylsalicylsyra (650 mg) 4,4 och tramadol (100 mg) 4,8. COX-2-hämmare kan alltså anses likvärdiga med i dag bästa tillgängliga alternativ för akut smärta. Vid dysmenorré var rofecoxib likvärdigt med naproxen.

Båda preparaten har en effekt utöver placebo vid behandling av knäledsartros. Vid reumatoid artrit har de också en effekt jämfört med placebo och av samma storlek som naproxen respektive diklofenak. Notabelt är att hela 30

Om forskning kring diagnostiska test med karpaltunnelsyndrom som exempel

Sannolikhetskvoter

– ett stort steg mot att förklara nyttan med ett test

När företrädare för evidensbaserad medicin började med systematiska genomgångar av randomiserade och kontrollerade studier av behandlingars effekt och biverkningar förskräcktes man snabbt av den dåliga kvaliteten i forskningen. Stora brister i design och genomförande samt alltför små patientmaterial ledde till stora svårigheter att kunna göra en syntes av underlaget och avge ett tillförlitligt utlåtande avseende nyttan med behandlingen.

Än värre visade sig situationen vara inom forskning kring diagnostiska test, dvs allt från anamnestiska och kliniska fynd till röntgenundersökningar och laboratorieprover. Sällan är testet jämfört oberoende och blindat med en tillförlitlig referens (golden standard), och än värre är att patienter med redan påvisad sjukdom jämförs med helt friska kontroller. Mycket sällan innefattar sådana studier samtliga patienter som sökt med ett specifikt symtom (t ex svullet underben, domningar i handen, smärtor i bröstet), och oftast är materialet insamlat där testet görs (vid MR-kameran, på fysiologlaboratoriet eller hos handkirurgen) i stället för där patienten presenterar sina symtom första gången (oftast i primärvården).

Det är också uppenbart att testens prestanda inte kan presenteras enbart i form av sensitivitet och specificitet vars kliniska innebörd de flesta (liksom jag själv) har mycket svårt att hålla reda på. Införande av sannolikhetskvoter (likelihood ratios, LR), är därför ett stort steg

mot att förklara nyttan med ett test. Ett LR-värde som är mycket större än ett visar att ett »positivt» testutfall påtagligt ökar sannolikheten för att sjukdomen föreligger. Vid ett »negativt» testutfall talar ett LR-värde långt mindre än ett starkt emot diagnosen. Ett LR-värde kring ett visar att testet inte bidrar till att föra den diagnostiska processen framåt och testet kan undvaras.

De senaste åren har ett empiriskt underlag växt fram, bl a med hjälp av Cochranesamarbetet, för hur diagnostiska test skall utvärderas och tolkas. Under hösten kommer Läkartidningen, i serien om evidensbaserad medicin, att reda ut dessa begrepp ytterligare.

Förekomsten av karpaltunnelsyndrom i den vuxna befolkningen är någon procent, men betydligt högre hos kvinnor samt vid diabetes, graviditet eller RA. För långdragna symtom har kirurgi tillgripits men också lokala steroidinjektioner har visats effektiva.

I JAMAs serie The rational clinical examination publicerades nyligen en systematisk översikt över vilken nytta man har av anamnes och status vid undersökning av en patient med domningar, stickningar eller smärta i en hand. En extensiv litteratursökning identifierade endast tolv artiklar med acceptabel metod och kvalitet. Elektrofysiologisk undersökning är i dag det slutliga diagnostiska testet men utgör egentligen ingen idealiskt »golden standard».

Bäst förmåga att bekräfta diagnosen

karpaltunnelsyndrom hade nedsatt känslighet för nålstick på pekfingerets palmarsida när man jämförde mot lillfingeret (LR 3,1), liksom ett typiskt mönster när patienten fick teckna utbredning av domning, smärta, parestesier och känselnedsättning på ett Katz handdiagram (LR 2,4). Fyndet av nedsatt styrka i abduktion av tummen bidrog också till att styrka diagnosen (LR 1,8). Om patienten ritade ett Katzdiagram som klassades som osannolikt eller uppvisade normal styrka i tumabduktion så kunde man med stor sannolikhet utesluta karpaltunnelsyndrom (LR 0,2 respektive 0,5). Patientens ålder, förekomst av bilaterala eller nattliga symtom, tenaratrofi, nedsatt sensibilitet med 2PD, stämgaffel eller monofilament bidrog inte till diagnostiken. Inte heller det klassiska Tinels tecken har klinikern någon nytta av.

Således bör klinikern som möter patient med domningar i handen i första hand undersöka sensibilitet för stick, styrka i tumabduktion och be patienten rita en symtomkarta. En reservation är på sin plats då en idealisk »golden standard» för karpaltunnelsyndrom saknas och då inga data härrör från ett oselekerat primärvårdsmaterial av patienter.

JAMA 2000; 283: 3110-7.

jama.ama-assn.org/issues/v283n23/full/jrc90001.html

procent var »responders» på placebo i dessa studier medan andelen bara ökade till 45–50 procent vid aktiv behandling. Att COX-2-hämmarna har lägre frekvens endoskopiska erosioner är väl belagt men får anses som ett surrogat-effektmått. Dessutom har i några studier doser av ibuprofen på 2 400 mg använts som jämförelse, vilket sällan är en kliniskt använd dos. Faktum är dock att COX-hämmarna inte skiljde sig från placebo vad gäller erosioner när tre studier och 1 500 patienter poolats.

De kliniskt viktigaste biverkningarna är självklart perforationer, ulkus och blödning (PUB), vilka genomgående använts som ett aggregerat biverkningsmått. En metaanalys gjordes av drygt 5 000 patienter, behandlade med rofe-

coxib (25 mg) placebokontrollerat fyra månader och följda med strikta diagnostiska kriterier. Frekvensen av PUB skilde sig inte mellan rofecoxib och placebo. Rofecoxidbehandlade patienter hade 49 procent (95-procentiga konfidensintervall 0–74 procent) lägre relativ risk för PUB än de som fått NSAID. Svagheten i denna analys är att den baseras på endast 38 allvarliga komplikationer.

Till slut handlar det om kostnader och risker. En amerikansk analys skattade kostnaden hos en lågriskpatient (risk 0,4 procent) för att förhindra en PUB genom att generellt använda COX-2-hämmare i stället för NSAID, till 400 000 dollar. Om det i stället rör sig om en patient över 75 år eller med ti-

digare ulkus stiger risken till 5 procent med NSAID och kostnaden för att förhindra en perforation, ulkus eller blödning sjunker till 30 000 dollar. I översikten redovisas också viktiga riskanalyser utifrån populationsperspektiv. Medan vi inväntar utlovade medicinska metodutvärderingar och hälsoekonomiska analyser av preparaten så rekommenderas läkemedelskommittéer och alla med intresse för rationell läkemedelsterapi att läsa Bandoliers viktiga och noggranna genomgång.

Cox-2 roundup. Bandolier 2000; 75

www.jr2.ox.ac.uk/bandolier/band75/b75-2.html

Två systematiska översikter inom antikonceptionsområdet

Subkutana implantat och hormonspiral – jämförelser

De brittiska hälsovårdsmyndigheterna, NHS, satsar ambitiöst på uppbyggnaden av systematiska kunskapsunderlag som för närvarande publiceras i snabb takt. Den senaste rapporten innehåller två systematiska översikter om implanterbara gestagena antikonceptionsmetoder; subkutana implantat (Norplant, Implanon) respektive sk hormonspiral (Levonova) i syfte att värdera dessa metoder i jämförelse med andra reversibla antikonceptionsmetoder. Som primärt effektmått använder forskarna inträffade graviditeter respektive benägenhet att fortsätta med metoden.

Granskningen av litteraturen om subkutana implantat bygger på 34 jämförande studier. Utvärderingen visade

ingen signifikant skillnad i graviditetsfrekvens mellan de olika metoderna. Sannolikheten att fortsätta var ca två gånger större med implantat än med metoder som p-piller och p-spruta.

För hormonspiral bygger granskningen på 29 interventionsstudier. I samtliga dessa jämförs olika gestagenimpregnerade spiraler med varandra eller med kopparspiral. Man fann ingen skillnad mellan hormonspiraler och kopparspiraler, innehållande >250 mm³ koppar, avseende oönskade graviditeter. Det var dock signifikant mindre risk för graviditet med hormonspiral jämfört med kopparspiral innehållande <250 mm³ koppar.

Däremot var det färre kvinnor som fortsatte med hormonspiral i mer än 5 år

jämfört med kvinnor som använde kopparspiral med 250 mm³ koppar.

Rapporten är en gedigen genomgång av området. Även om en betydande del av underlaget utgörs av studier från andra miljöer än vår, kan den vara intressant för svenska förhållanden genom den kunskap om en värdefull princip för antikonception som lyfts fram. Betydande svenska forskningsinsatser har ju också gjorts inom detta område. Rapporten kan rekommenderas för läkare och barnmorskor verksamma inom antikonceptionsområdet.

Kerstin Nilsson

kerstin.nilsson@orebroll.se

Health Technology Assessment 2000; 4(7).
<http://www.hta.nhsweb.nhs.uk/execsumm/SUMM407.HTM>

Inget stöd i valet mellan peroralt och intravenöst penicillin

I den kliniska vardagens alla små beslut står vi påfallande ofta utan mer påtagligt vetenskapligt underlag (se LT 36/00). En process, som nyttjar den evidensbaserade medicinens grunder med fokuserad fråga, strukturerad sökning och granskning samt värdering i relation till vår vardag, erbjuder verktyg för att ta sig an sådana frågeställningar.

Den danska »SBU-enheten» (MTV) beskriver i en kort skrift hur infektionsläkarna i Odense angrep frågan om huruvida vuxna patienter, utan bakomliggande åkommor som läggs in med pneumoni, skall behandlas med penicillin peroralt eller intravenöst. Processen

är väl beskriven och så är även det nedslående resultatet att endast en randomiserad undersökning studerat kliniska utfall.

Man fann inga skillnader i mortalitet, behov av antibiotikaskifte eller antal sängdagar mellan de två administrationsätten. I två studier var peroral behandling förenad med kortare vårdtid om man startade med kort intravenös behandling för att därefter skifta till tabletter, jämfört med längre tids intravenös tillförsel. Organisatoriskt var infusionsbehandlingen förenad med betydligt större tidsåtgång och kostnader. Den största ekonomiska besparingen

uppkom dock om inläggningstiden kunde reduceras.

Trots att frågan är dåligt undersökt talar mycket för en större frikostighet med peroral behandling även för inläggande patienter. Intressant att notera är att den danska infektionskliniken använde penicillin och inte cefalosporin, vilket annars tycks ha blivit något av ett förstahandsmedel även när en patient läggs in med samhällsförvärd pneumoni.

Medicinsk problemställning i MTV-perspektiv. Penicillinbehandling ved pneumoni: oralt eller intravenöst? Sundhedsstyrelsen, Köpenhamn, 2000.
www.dihta.dk

Var tredje hudbehandling har stöd i randomiserade studier

Framväxten av evidensbaserad medicin har lett till försök att, i studier med god vetenskaplig design, värdera hur stor andel av våra behandlingar som egentligen har visad effekt.

Svenska och brittiska undersökningar har visat att drygt hälften av de behandlingsåtgärder (t ex farmaka, kirurgi eller psykosociala insatser) som använts inom invärtes medicin och psykiatri är baserade på randomiserade, kontrollerade studier, vilka (tillsammans med välgjorda systematiska översikter) anses ha det högsta bevisvärdet. Inom

brittisk allmänmedicin och kirurgi var denna andel cirka en tredjedel.

Danska dermatologer studerade ett randomiserat urval av patientjournaler från Bispebjergs sjukhus och noterade den definitiva diagnosen samt vilken huvudsaklig behandling som gavs. Om behandlingen var visat effektiv i någon randomiserad studie bedömdes evidensen som god. Som en lägre bevisnivå accepterades en visad behandlingseffekt extrapolerad från en diagnos till en närliggande diagnos, t ex från seborroisk till atopisk dermatit.

Hos 115 patienter fann man 62 olika diagnoser och gav 118 olika behandlingar. I 38 procent fanns god evidens för nyttan med behandlingen, i 33 procent sämre stöd och 23 fanns inget stöd alls. Resterande behandlades med expectans. Således tycks dermatologi väl matcha andra specialiteter avseende bristen på god evidens för de dagliga behandlingarna.

Ugeskr Læger 2000; 162: 4382-5.

Svensk frontlinjeforskning

Bäst nytta av strålbehandling vid tidig bröstcancer hos unga kvinnor

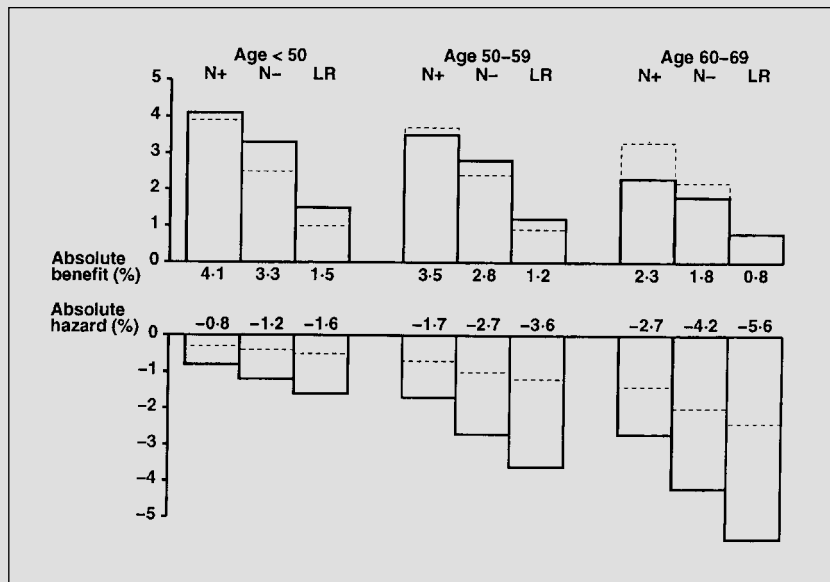
Inom behandlingen av bröstcancer har det under många år rått en tveksamhet om radioterapi efter kirurgi förbättrar överlevnaden eller ej. Det finns 40 väl genomförda randomiserade studier varav 34 studerat effekten efter mastektomi och sex efter bröstbevarande kirurgi. Samtliga studier visar att postoperativ strålbehandling effektivt förebygger lokala recidiv i operationsområdet men bara tre visar att överlevnaden förbättras.

Före sommaren publicerade Lancet en välgjord metaanalys av dessa studier, varav flera är svenska. Analysen innehåller individuella patientdata från 20 000 kvinnor behandlade med mastektomi (16 000) eller bröstbevarande kirurgi (4 000), som postoperativt randomiserats till antingen strålbehandling eller enbart kontroll. Analysen inkluderade endast patienter med »tidig» bröstcancer, dvs begränsad till bröstet eller lokala körtlar (körtelpositiva).

Uppföljning har skett upp till 20 år efter behandlingen. Under de första 10 åren reducerade strålbehandlingen risken för lokalt recidiv från 27 till 9 procent, en absolut riskreduktion på 18 procentenheter. Mellan år 10 och 20 skedde ytterligare en liten reduktion så att den slutliga skillnaden avseende lokala recidiv mellan grupperna var cirka 20 procentenheter.

När det gäller den mer kontroversiella frågan om överlevnad observerades efter 10 år en överlevnad hos de strålbehandlade patienterna som var 2,3 procentenheter större än hos dem som inte fått strålning ($P=0,06$). Efter 20 år hade effekten minskat till 1,2 procentenheter.

Uppdelat på dödsorsak visade det sig att det fanns en signifikant förbättrad överlevnad i bröstcancer hos de strålbehandlade patienterna på 3 procent efter 10 år och 4,8 procent efter 20



Absolut nytta av och risk med postoperativ strålbehandling. Effekt på överlevnad upp till 10 år (streckad linje) och 20 år (heldragen linje) om den proportionella förändringen i bröstcancerrelaterad dödlighet och icke bröstcancerrelaterad dödlighet appliceras på olika typer av patienter (13,2 procents reduktion i bröstcancerrelaterad dödlighet och 21,2 procent ökning i icke bröstcancerrelaterad dödlighet. Båda efter 2 år). Figuren är hämtad ur författarens arbete (Lancet 2000; 355: 1757-70) och publiceras här med tillstånd av The Lancet Ltd.

N+ = patienter med axillära lymfkörtelmetastaser.

N- = patienter utan axillära lymfkörtelmetastaser.

LR = patienter med låg risk för lokalt recidiv och utan axillära lymfkörtelmetastaser.

år. Denna effekt motverkades av en signifikant överdödlighet i andra sjukdomar, framför allt hjärt-kärlsjukdom, på 1 procent efter 10 och 4,3 procent efter 20 år.

Det stora antalet patienter möjliggjorde en analys av effekter och bieffekter i olika subgrupper (se figur). Av denna framgår att de kvinnor som har störst nytta av postoperativ strålbehandling är unga kvinnor med tumörsjukdom spridd till armhålan. Hos kvinnor över 59 års ålder överväger de negativa effekterna av strålbehandling. Vidare framgår att strålbehandlingens positiva effekter kommer ti-

digt, från 2 till 10 år, medan bieffekter på hjärta och kärl kommer senare, efter 10 år.

Resultaten har i Sverige medfört att vårdprogram delvis ändrats så att fler kvinnor än tidigare får strålbehandling mot regionala lymfkörtelstationer. Dessa resultat har varit kända sedan 1995 och en ny metaanalys görs för närvarande i Oxford och kommer att presenteras senare i höst. Då föreligger ytterligare 5 års uppföljning.

Göran Liljegren
goran.liljegren@orebroll.se

Lancet 2000; 355: 1757-70.

Skriv till oss!

Läkartidningen vill uppmuntra läsarna att själva skriva kommentarer och referat av systematiska översikter med relevans för en bredare läsekrets. Uppslaget »Nya rön» tar upp evidensbaserad medicin med betoning på referat av bra systematiska översikter och metaanalyser. Andra aktuella artiklar som berör evidensbaserad medicin kan också vara av intresse för tidningens läsare.

Om du vill skriva Kontakta Mats Eliasson, medicinsk redaktör.

Mats.Eliasson@nll.se
eller 070-513 02 93.

Myten om kunskapsexplosionen – rökavvänjning åskådligt exempel

Det är vanligt att påstå att den vetenskapliga utvecklingen är »explosionsartad» vilket skulle ligga bakom läkarnas »oförmåga att hänga med i utvecklingen». Ingen kan ifrågasätta att det publiceras mycket medicinsk vetenskap men är det liktydigt med att nya, viktiga och relevanta kunskapstillkott duggar tätt i den kliniska vardagen?

Jag har upprepade gånger konfronterat specialister och allmänläkare med frågan om vilka för patienten viktiga nyheter som tillkommit de senaste åren. Har de nya metoderna lett till att gamla behandlingar ersatts av nya och varit patienterna till bättre nytta än vad sjukvården tidigare kunnat åstadkomma? Oftast kan man redovisa en handfull nya behandlingar eller diagnostiska metoder, men knappast en störtskur.

Efter ett årtiondes arbete med systematiska medicinska utvärderingar blir det allt mer uppenbart att det i de flesta specialiteter går lång tid mellan de stora nytillskotten i den terapeutiska eller diagnostiska arsenalen. Mycket handlar om marginellt förbättrade effekter eller diagnostiska prestanda.

Till exempel kan nämnas att SBUs projektgrupper återsamlas tre år efter avslutade rapporter för att ta ställning till om så mycket nytt tillkommit att revision av rapport och slutsatser är nödvändig. SBUs ryggrapport från 1991 ersattes av en ny version först i år, hypertoni-rapporten från 1995 kommer att revideras under nästa år och rapporten om bentäthetsmätning från samma år upgraderades nyligen i Läkartidningen med tolkningen att slutsatserna var väsentligen oförändrade.

För två år sedan publicerades SBUs rapport Metoder för rökavvänjning (baserad på litteratursökning fram till sommaren 1997) som framhöll nyttan av kort rådgivning vid alla vårdens kontakter med rökare samt rekommenderade användning av nikotinersättningsmedel.

Två nya och oberoende stora systematiska översikter om rökavvänjning presenterades nyligen. De amerikanska hälsovårdsmyndigheterna publicerade sina kliniska riktlinjer, vars åtta slutsatser är närmast identiska med SBU-rapportens förutom att man även finner att

bupropion (som 1998 inte var godkänt i Sverige) ökar sannolikheten för rökfrihet.

Cochranesamarbetet redovisar en sammanställning av de 20 systematiska översikter som dess Tobacco Addiction-grupp skrivit. Även här betonas nyttan av kort rådgivning och nikotinersättning. Man framhåller att bupropion hitintills bara studerats i fyra studier, vilket ska jämföras med de drygt 70-talet studier där nikotinersättning undersökts, och att en adekvat jämförelse mellan dessa båda behandlingar för närvarande inte är möjlig.

Sammanfattningsvis är det uppenbart att när väl genomförda systematiska och kritiska utvärderingar av medicinska insatser föreligger, tar det oftast lång tid innan deras slutsatser och därmed klinisk praxis, behöver förändras.

BMJ 2000; 321: 355-8.

<http://www.bmj.com/cgi/content/full/321/7257/355>

JAMA 2000; 283: 3244-54.

<http://jama.ama-assn.org/issues/v283n24/abs/jst00005.html>

Symposier sponsrade av industrin tar över vid vetenskapliga möten

En aktuell ledare i Lancet identifierar flera viktiga problem i relationen mellan läkemedelsindustri, forskare och kliniker som riskerar att leda till beslut om behandling som inte alltid är baserade på bästa vetenskap. I flera artiklar har nyligen påpekats att industrin ofta väljer att inte jämföra sina nya preparat med bästa tillgängliga standardbehandling. I stället väljer man medvetet att behandla kontrollgrupperna med medel med sämre effekt. Detta får självklart till följd att deras egna preparat verkar medföra större nytta för patienterna än vad som verkligen är fallet. Exempel från såväl studier av myelom som inom kardiologi redovisas.

Richard Horton, Lancets chefredaktör, påpekar dessutom hur industrisponsrade satellitsymposier breder ut sig vid vetenskapliga möten. Han exemplifierar från årets europeiska kardiologmöte (ESC) och frågar sig om det finns anledning till oro, när knappt 20 procent av det årliga utbildningsinnehållet är kommersiellt understött och när denna andel ökar?

Tidigare studier har visat att publikationer från industrisymposier, som tid-

skriftssupplement, ofta hade missledande rubriker, använde varunamn i stället för generiska substansnamn och i mindre utsträckning hade genomgått »peer review». Signifikant fler artiklar i industrisponsrade supplement än i fristående publikationer »bevisade» att läkemedlet i fråga hade god effekt, 98 respektive 79 procent (P=0,01). Vid ESC förekom också missbruk av föreningens logotyp på vilseledande marknadsföringsmaterial från Bayer.

Ledaren betonar att vetenskapliga sällskap måste inse att deras goda rykte är beroende av såväl verkligt som av omgivningen uppfattat oberoende. En förlust av detta oberoende kan visa sig svår att återhämta sig från. Vi har anledning att återkomma till och ytterligare skärskåda i vilken mån kommersiella intressen står emellan klinikerna och deras möjlighet att bibringa patienten vård baserad på oberoende granskning och sammanställning av vetenskap.

Lancet 2000; 356: 959-60.

<http://www.thelancet.com/newlancet/reg/issues/vol356no9234/commentary959.html>

Ann Intern Med 1996; 124: 485-9.

Seminarium om Cochranesamarbetet

Många svenska läkare och forskare har hört av sig för att ta del av hur man konkret kan delta i och bidra till Cochranesamarbetet. Det nordiska Cochranecentret, i samarbete med SBU, vill därför bjuda in alla intresserade till ett seminarium i Stockholm.

Eftermiddagen onsdagen den 7 februari kommer Peter C Götzsche från Köpenhamn för att introducera verksamheten.

Svenska redaktörer från flera olika review groups inom Cochranesamarbetet deltar och berättar om hur det konkreta arbetet går till.

Vi kommer också att diskutera hur vi kan sprida kunskap om Cochrane i Sverige, t ex inom våra specialistföreningar och fakulteter.

Avsätt gärna tid för mötet redan nu och meddela mig ditt intresse.

Definitiv inbjudan kommer senare i Läkartidningen.

Mats Eliasson
Mats.Eliasson@nll.se

Irritable bowel syndrome – fortsatt svårbehandlat

En aktuell genomgång av samtliga studier där nyttan av farmakologisk behandling av irritable bowel syndrome (IBS) utvärderats mellan 1966 och 1999, visade att endast 70 av 283 studier var av godtagbar kvalitet [1]. Kraven på studierna var i vissa avseenden högst modesta: minst två veckors behandling av fler än 10 patienter. Översikten redovisar detaljerat de 28 mest välgjorda studierna, och summerar resultaten av de övriga 42.

Bulkmedel visar sig ha dålig effekt på den buksmärta som vanligtvis är det mest invalidiserande symtomet. Muskelrelaxerande läkemedel däremot är väldokumenterade mot smärta och detta gäller samtliga fyra substanser som utvärderats i de bästa studierna.

Tyvär är inget av dessa läkemedel registrerat i Sverige.

I gruppen prokinetiska läkemedel återfinns cisaprid, som dock inte har någon övertygande effekt. Om det dominerande symtomet är diarré är däremot loperamid väldokumenterat som terapeutikum. Tricykliska antidepressiva anses otillräckligt utvärderade. Endast en studie avseende amitriptylin är av hög kvalitet, och den effekt som ofta påvisats kan inte säkert skiljas från den antidepressiva verkan. SSRI-preparat har inte utvärderats vid IBS.

Alosetron är, en i Sverige inte godkänd, 5-HT₃-receptorantagonist. En av översiktens senaste studier, som undersökt alosetron i olika doser och i jämförelse med placebo, visar att alosetron

hjälpes mot smärta och diarré, men bara hos kvinnor [2]. I november 1999 publicerades ännu en studie av alosetron, denna gång i jämförelse med ett i Sverige icke registrerat kramplösande läkemedel, mebeverine, och denna gång deltog bara kvinnliga patienter [3]. Studien genomfördes med deltagare även från 11 svenska centra, och visade att alosetron var effektivare än mebeverine.

Anders Rönnblom

anders.ronnblom@medsci.uu.se

1. *Ann Intern Med* 2000 Jul 18; 133: 136-47.

(<http://www.annals.org/issues/v133n2/full/200007180-00013.html>)

2. *Aliment Pharmacol Ther* 1999; 13: 1149-59.

3. *Aliment Pharmacol Ther* 1999; 13: 1419-27.

Lågdos p-piller ökar risken för ischemiskt slaganfall – men absoluta risken är mycket låg

Systematiska översikter med åtföljande metaanalyser är inte bara värdefulla verktyg avseende nyttan av behandlingar utan även för att värdera risker och biverkningar. För att bedöma behandlingseffekter utnyttjar vi oftast randomiserade och kontrollerade studier men för att identifiera ovanliga biverkningar måste uppgifter sammanställas från observationsstudier, antingen i form av fall-kontrollstudier eller uppföljningar av stora kohorter. Bristen på randomisering i denna typ av studier leder till stora risker för systematiska fel (»bias«).

Vanligaste sättet att minska bias är att statistiskt korrigera för olikheter mellan grupperna avseende viktiga faktorer som påverkar det studerade utfallet, men sådana exerciser kan aldrig ga-

ranterera att de systematiska felen undanröjs. Problemet med observationsstudier kommer att beröras i en kommande Medicinsk kommentar.

Sambandet mellan ischemiskt slaganfall och p-piller har debatterats under många år. Eftersom risken för slaganfall i de aktuella grupperna är mycket låg kommer frågan aldrig att kunna lösas i en randomiserad studie. Enda utvägen är att klokt använda observationella data.

Amerikanska forskare publicerade nyligen en mycket välgjord systematisk översikt baserad på en extensiv litteratursökning med strikta inklusionskriterier vilket resulterade i 16 studier, varav tre var fall-kontrollstudier och 13 observationsstudier. Pågående bruk av p-

piller ökade risken för ischemiskt slaganfall. Efter korrektion för rökvanor och hypertoni fann man en dosberoende effekt, men även lågdos östrogen (<50 mg) ledde till en fördubblad risk (95-procentigt konfidensintervall 35–174 procent).

Den evidensbaserade medicinen vill dock framhäva absoluta risker och bilden blir då något annorlunda. Nyttjande av lågdos p-piller hos icke-rökande, normotensiva kvinnor leder till ett extra insjuknande i ischemiskt stroke per 24 000 som använder preparaten under ett år.

JAMA 2000; 284: 72-8.

<http://jama.ama-assn.org/issues/v284n1/abs/jrv90079.html>

Slutligen några ord om Internetadresser ...

På de här sidorna anger vi oftast webbadresser där läsaren själv kan ta del av de dokument vi redovisar. Cochrane-översikter är tillgängliga via www.ronden.se för registrerade användare och alla rapporter från medicinska utvärderingsorganisationer (HTA-rapporter) brukar vara fritt åtkomliga i fulltext på nätet.

De medicinska tidskrifterna har varierande policy som också förändras över tiden. Många tidskrifter har sina innehållsförteckningar utlagda på nätet och

dessas kan man prenumerera på via e-post. Enstaka tidskrifter medger allmän tillgång till alla artiklar, antingen i textformat eller, ännu bättre, i pdf-format (som är läsbara med gratisprogrammet Acrobat Reader). Använd pdf-formatet för att spara, läsa och skriva ut artiklarna med samma utseende som den tryckta tidskriftssidan.

En del tidskrifter nöjer sig med att man registrerar sig för att få full eller partiell tillgång till artikelarkivet, men många kräver ett lösenord kopplat till

en prenumeration på pappersupplagan av tidskriften. Det innebär att om kliniken har en prenumeration, som omfattar även onlineversionen, får man tillgång till tidskriften i fulltext via Internet.

Ditt lokala medicinska bibliotek kan ha en prenumeration och därmed också lösenord. Universitetsbiblioteken i Sverige har tillsammans upphandlat tidskrifter från vissa stora förlag genom ett nationellt avtal. Tidskrifter från dessa förlag finns då tillgängliga i fulltext via universitetets nätverk.

Kliniska beslutsregler kan förenkla diagnostik – exemplet fotledstrauma

Trots att vår vardag är fylld med diagnostiska överväganden har denna process väckt betydligt mindre intresse än olika aspekter på patienters behandling. Detta är anmärkningsvärt med tanke på att den bäst bevisade behandling kan vara till ingen nytta eller rent av skadlig om diagnosen är fel. Tilltagande sofistikerad av röntgenologiska och biokemiska metoder har skapat en falsk illusion om att klinisk diagnos är mindre väsentlig än förr. Tvärtom kan man hävda att bra diagnostiska metoder bör användas i situationer där man genom god klinisk diagnostik sorterat fram patienter som med större sannolikhet har den misstänkta sjukdomen. Detta ger bättre diagnostiska skärpa. Därför behövs bra instrument för att sälla fram vilka patienter som ska gå vidare till ytterligare utredning som kan vara både farlig eller dyrbar. Risken för oanade konsekvenser av bifynd eller falskt positiva fynd bör också övervägas.

Vår kliniska erfarenhet förser oss med en intuitiv känsla för vilka fynd i patientens sjukhistoria eller status som bör väcka misstanke om eller bekräfta en diagnos. Ofta är denna förmåga remarkabelt träffsäker men inte alltid. En klinisk beslutsregel kan ses som ett kliniskt verktyg som kvantifierar i hur stor utsträckning individuella komponenter i anamnes, status och basala laboratorieprov bidrar till rätt diagnos, prognos eller prediktion av behandlingssvar. Beslutsregler ger en formaliserad möjlighet att testa, förenkla och skärpa den

diagnostiska förmågan. Beslutsregler kan vara till mest nytta i komplicerade beslutsituationer som t ex innebär risktaganden eller krav på besparingar utan att detta försämrar vårdkvaliteten. Beslutsregler för bl a diagnostik av halsinfektioner och lungemboli, screening för alkoholism och depression samt riskbedömning vid antikoagulantabehandling finns publicerade på följande hemsida: (<http://med.mssm.edu/ebm>).

Med utgångspunkt i det kliniska problemet fotledstrauma och ställningstagande till röntgen beskrivs i JAMA hur en klinisk beslutsregel skapas och valideras. I ett första steg identifieras, med hjälp av logistisk regression, faktorer i status och anamnes som hos en grupp patienter hade ett högt prediktivt värde för diagnosen i fråga.

I ett andra steg testas dessa faktorer i nya patientgrupper och kliniska miljöer för att bekräfta att de ursprungliga resultaten inte berodde på slump, speciella patientpopulationer eller speciellt duktiga kliniker.

Därefter värderas i vilken mån implementering av beslutsregeln verkligen leder till förbättrad handläggning och till den slutliga målsättningen: bättre effekt för patienten och/eller minskade kostnader. Dessa tre steg utgör också en enkel evidensgradering av hur väl underbyggd en klinisk beslutsregel är. Självfallet bör vi i första hand ta till oss regler som testats och validerats hela vägen fram till bevisad patientnytta.

De s k Ottawa-reglerna föreslår att

röntgen vid fotledsskada endast är nödvändig om patienten har smärta i en väl definierad »malleolär zon» och palpationsömhets på den posteriora delen eller spetsen av malleolen. Även smärtor kombinerade med oförmåga att belasta omedelbart efter traumat samt vid ankomsten till akutmottagningen ska leda till röntgenundersökning.

Ottawareglerna utvecklades vid två stora universitetssjukhus och bekräftades senare i ett stort antal andra kliniska miljöer. Diagnostiska prestanda är goda. Alla patienter med fraktur urskiljs (sensitivitet 100 procent, 95 procent konfidensintervall, 93–100). Regeln urskiljer 40 procent av dem som inte har fraktur och användandet av regeln skulle alltså kunna minska antalet röntgenundersökningar med 40 procent. Implementeringsundersökningar visar en påtaglig reduktion av antalet beställda undersökningar och endast tre missade frakturer av 906 patienter handlagda med Ottawa-reglerna. Ingen av dessa krävde dock behandling.

Kliniska beslutsregler rymmer en stor potential genom introduktion av ett mer evidensbaserat tänkande inom diagnostiken som kan leda till bättre utnyttjade resurser utan försämring av patientens vård.

JAMA 2000; 284: 79-84.
<http://jama.ama-assn.org/issues/v284n1/rfull/jml90005.html>

Sängvila – en potentiellt riskabel behandling i behov av utvärdering

Redan SBU:s ryggrapport från 1991 pekade på farorna med sängvila vid akut ryggonn vilket bekräftades när rapporten uppdaterades i år. Australiska forskare redovisar nu en systematisk översikt om nyttan av sängvila jämfört med tidig mobilisering vid sjukdom och diagnostik. Man identifierade 39 studier vid 15 olika medicinska tillstånd, omfattande knappt 6 000 patienter. Studierna var mycket heterogena vad gäller uppläggnings av sängvila, kontrollbehandling och uppföljningstider, varför man avstod från att genomföra någon metaanalys.

Effekten av sängvila vid olika ingrepp eller diagnostiska åtgärder var i sju studier bättre än i kontrollgruppen

som mobiliserades tidigt, men i inget fall var skillnaden signifikant. Å andra sidan utföll 26 studier effektmått sämre med sängvila och i nio fall var den skillnaden signifikant. Sängvila som primär behandling vid sjukdom uppvisade likartade resultat med 25 av 32 resultat sämre vid sängvila varav hela nio skillnader var statistiskt signifikanta. Många av de undersökta sjukdomarna handläggs dock i dag med helt andra principer varför fynden mest har historiskt intresse.

Hur patient ska handläggas efter lumbalpunktion har däremot fortfarande stor relevans på neurolog- och akutmottagningar. Vid årets stora Cochrane-möte presenterade en österrikisk grupp

en systematisk sammanställning av 15 studier med 2 000 patienter som efter lumbalpunktion randomiserades till antingen sängvila eller snabb uppstigning. Frekvensen huvudvärk i efterförloppet uppvisade stor spridning, mellan 9 och 60 procent. Inga studier visade att sängvila minskade risken för huvudvärk och i två studier ökade till och med huvudvärksfrekvensen med vila. Det finns alltså inte mycket som talar för en läkande kraft av att ligga till sängs vid sjukdom eller efter olika ingrepp.

Lancet 1999; 354: 1229-33.
Muellner M. 8th International Cochrane Colloquium 2000: 125.
www.cochrane.org

Snabbt omhändertagande av traumapatienter är och förblir viktigast

Av dem som avlider vid trauma dör hälften på skadeplatsen och 20 procent till följd av komplikationer. De resterande 30 procenten dör inom några timmar. Det initiala omhändertagandet är avgörande för prognosen, därav begreppet »the golden hour». Hur skall denna förvaltas? Mot varandra står två skolor. Å ena sidan: enkla, livräddande åtgärder och därefter snabbast möjliga transport till sjukhus för definitiva åtgärder. Å den andra: avancerat prehospitalt omhändertagande, kanske även om det innebär fördröjd transport. I Sverige diskuteras frågan, och exempelvis kravet på sjuksköterskekompe-

tens för ambulanspersonal är en åtgärd som möjliggör mer avancerade åtgärder på skadeplatsen.

En metaanalys (174 artiklar i Med-Line, varav totalt 49 uppfyllde analyseringskriterierna men endast 15 analyserades avseende »outcome») pekar dock på en tydlig överdödlighet hos patienter som initialt omhändertas med avancerade åtgärder (intubation, perifer nål och vätskeinfusion, antichockbyxor) oddskvot 2,92 (2,59 efter justering för studiekvalitet, metodologi, patientgrupper). Skadepanoramata är enormt, och att jämföra populationer och subgrupper är svårt. Många andra faktorer

spelar in: typ av trauma, typ av och tid för transporter m m. Alternativa metoder, som intraosseös infusion och hypertont koksalt, fanns ej representerade i analysen.

Att applicera (även välgrundade) statistiska resultat på enskilda patienter förblir svårt och är inte att rekommendera, men studien inskräper ändå ett viktigt budskap: dyrbar tid får inte gå till spillo på onödiga åtgärder!

Christian Olsson

christian.olsson@thorax.uas.lul.se

J Trauma 2000; 49: 584-99.
www.jtrauma.com

Primärpreventiv läkemedelsbehandling av hyperkolesterolemi – en ny metaanalys

Hyperkolesterolemi är en väldokumenterad riskfaktor för hjärtkärlsjukdom där enskilda primär- och sekundärpreventiva studier bidragit med kunskaper om möjliga behandlingsvinster vid intervention.

I en ny metaanalys har man undersökt effekter av primärpreventiv läkemedelsbehandling riktad mot hyperkolesterolemi. Urvalskriterierna var randomiserade, placebokontrollerade studier av minst ett års duration, med målsättning att påverka morbiditet och mortalitet, samt publicerade på engelska. Det blev sammanlagt fyra stycken (LRC, HHS, WOSCOPS, AFCAPS/TEXCAPS), varav bara den sistnämnda innehöll kvinnor. Totalt ingick endast cirka 900 kvinnor av de knappt

20 000 personer som översikten redovisar. Behandling med lipidsänkande läkemedel kunde signifikant minska antalet nya kardiovaskulära händelser. Koronar morbiditet minskade, oddskvot 0,70 (95 procents konfidensintervall: 0,62–0,79) liksom koronar död, oddskvot 0,71 (0,56–0,91). Total mortalitet minskade däremot inte signifikant, oddskvot 0,94 (0,81–1,09). Resultaten blev något mer gynnsamma för koronar morbiditet och mortalitet om endast statinbaserade studier medräknades, men total mortalitet ändrades ej.

Författarna drar slutsatsen att primärpreventiv behandling möjligen kan minska totalmortalitet ifall patientgrupper med högre absolut kardiovaskulär risk väljs ut för behandling. De absolu-

ta riskerna var genomgående mycket låga, t ex minskade koronar mortalitet från 1,5 till 1,1 procent med behandlingen. Tiden för uppföljning bör också beaktas eftersom de här använda studierna endast hade en medeluppföljningstid av 5–7 år. Slutsatserna i översikten stödjer i allt väsentligt vad man kom fram till i Läkemedelverkets workshop om behandling av höga blodfetter (1999), särskilt i avsnittet med kliniska och hälsoekonomiska aspekter på värdet av och begränsningarna med primärpreventiv behandling.

Peter Nilsson

peter.nilsson@medforsk.mas.lu.se

BMJ 2000; 321: 983-6.
http://www.bmj.com/cgi/content/full/321/7267/983

Kronisk prostatit – ingen evidens för någonting

Det finns idag inget diagnostiskt test som kan betraktas som »gold standard» för kronisk prostatit och de diagnostiska studier som finns håller låg kvalitet. De få kontrollerade behandlingsstudierna är metodologiskt svaga och studiegrupperna är små. Det finns idag ingen evidens för vare sig antibiotikabehandling eller behandling med α -blockerare.

Detta är slutsatserna i en strukturerad analys av den tillgängliga litteraturen om diagnostik och behandling av kronisk prostatit [1]. Författarna hade sökt efter kontrollerade diagnostiska studier och randomiserade, kontrollerade behandlingsstudier i MEDLINE, Cochrane library och i bibliografier i relevanta artiklar. Man fann 19 diagnos-

tiska studier och 14 behandlingsstudier som uppfyllde kriterierna. De diagnostiska artiklarna rörde diagnos av infektion och inflammation, immunologi, biokemi, ultraljud och psykologiska test. Inte någon av dessa diagnostiska principer kunde säkert användas för diagnosen kronisk prostatit. Behandlingsartiklarna rörde α -blockerare, finasterid, antiinflammatoriska medel, antibiotika, värmebehandling och andra varierande medikamenter. Det fanns inget vetenskapligt stöd för att någon av dessa behandlingar hade en gynnsam effekt för patienten.

Således är kunskapen om kronisk prostatit minst sagt dålig. En grundorsak till detta är säkerligen att diagnosen

som sådan är odefinierad och oklar. Två steg på vägen till bättring är troligen den nya klassifikation av prostatit som föreslagits av National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) [2] och det nya symtomformulär som tagits fram av National Institutes of Health [3]. Även om dessa instrument ännu inte är helt optimala så är de användbara verktyg för att börja nysätta upp diagnosen kronisk prostatit.

Jan Adolfsson

adolfsson@sbu.se

1. Ann Intern Med 2000; 133: 367-81.
2. JAMA 1999; 282: 236-37.
3. J Urol 1999; 162: 369-75.

Även lågdos-ASA ger magtarmblödningar vid långtidsprofylax

II Acetylsalicylsyra (ASA) har en utbredd användning som profylax mot hjärt-kärlsjukdomar, huvudsakligen i sekundärpreventivt syfte. En ökad risk för gastrointestinala blödningar har länge varit klarlagd vid långtidsbehandling med ASA, men en förhoppning har varit att lägre doser skulle minska dessa risker. En aktuell systematisk översikt från Oxford talar dock emot detta.

I en metaanalys av 24 randomiserade studier med 65 000 deltagare var risken för blödning 1,42 procent i placebogruppen och 2,47 procent i gruppen som fått ASA som profylax under i genomsnitt 28 månaders behandling. Detta ger en oddskvot på 1,68 (95-procentiga konfidensintervall 1,51 till 1,88). Risken med behandlingen kan också uttryckas som att »numbers needed to harm« var 106, dvs om drygt 100 patienter behandlas

med ASA under ca 2 år kommer en av dem att drabbas av en blödning från mag-tarmkanalen.

Studier, där man undersökt användningen av låga doser av ASA (50–162,5 mg), analyserades separat. I de åtta studierna med 50 000 deltagare var andelen med blödning närmast identisk med andelen i hela gruppen; 1,45 och 2,3 procent i placebo respektive i ASA-gruppen.

Risken var fortfarande signifikant förhöjd (oddskvot 1,59; KI 1,40; 1,81) om än något lägre än i högdosgruppen (oddskvot 1,96; KI 1,58; 2,43). Denna skillnad var inte signifikant (notera de överlappande konfidensintervallen) och inte heller i en sk metaregression var dosen av ASA signifikant relaterad till blödning.

I en åtföljande ledarkommentar med

titeln »Aspirin, like all other drugs, is a poison« understryks fyndet att inte ens lågdos-ASA är ofarligt. Detta leder till att man inte kan påvisa en, i sin helhet, positiv effekt med ASA-profylax för personer med liten risk för kardiovaskulära komplikationer, som t ex dem utan känd hjärtkärlsjukdom.

Ledaren pekar dock på att i översikter där sambandet ASA-dos och potentiellt livshotande blödningar studerats visas ett klart samband, åtminstone för doser över 300 mg. Dessa doser är å andra sidan inte aktuella vid profylaktisk behandling.

BMJ 2000; 321: 1183-7.

BMJ 2000; 321: 1170-1.

Megaöversikt belyser prevention och folkhälsoinsatser

II I samband med reformeringen av NHS (National Health Service) i England och Wales har en omfattande litteratursammanställning med titeln »Evidence from systematic reviews of research relevant to implementing the wider public health agenda« nyligen givits ut av NHS Centre for Reviews and Dissemination (CRD) vid University of York, England.

Dokumentet, som är skrivet av en arbetsgrupp ur Cochrane Collaboration, Campbell Collaboration och NHS, bygger på över 1 100 systematiska översikter och är främst ett evidensbaserat underlag för policybeslut men kan även fungera som stöd vid enskilda ställningstaganden. Den intresserade läkaren kan t ex via hemsidan (se adress nedan) snabbt få upplysningar om relevanta studier med hög evidens inom olika områden.

Rapporten är i princip ett källdokument innehållande korta sammanfattningar och referenser till systematiska översikter inom specifika områden men också till övergripande folkhälsoöversikter. Innehållet är uppdelat i »national contracts« med intentionen att reducera morbiditeten och mortaliteten inom följande områden: cancer, kranskärlssjukdomar och stroke, olycksfall samt mental ohälsa. Tillfört är dessutom kapitel om utbildning, social omvårdnad och social välfärd samt kriminalitet, droger och alkohol. Sökstrategier och referenser till samtliga citerade systematiska översikter är utförligt redovisade i två appendix.

Dokumentet är ett hjälpmedel vid besvarandet av följande frågor:

Vilka policies/strategier fungerar effektivt och bör prioriteras i folkhälsoarbetet?

Hur nås bäst de uppställda målen utifrån ovanstående prioriteringar? Vilka interventioner är lämpligast och hur genomförs dessa?

Vilket är behovet av ytterligare folkhälsoforskning? Inom vilka områden är behovet och vinsterna störst?

Denna brittiska rapport kommer kontinuerligt att uppdateras med nya systematiska översikter, allt för att förbättra förutsättningarna för folkhälsoarbetet.

Mats Halldin

(halldin@sbu.se)

Rapporten beställes via CRD,

e-postadress: crdpub@york.ac.uk

Priset är 20 pund (cirka 290 kr).

Hela dokumentet är dessutom tillgängligt på <http://www.york.ac.uk/inst/crd/wph.htm>

I korthet

Ingen gemensam patofysiologi vid whiplashskada.

II Varje år diagnostiseras i samband med trafikolyckor i Norge 2 000 nya fall av whiplashskada. Ökande incidens och tveksamheter avseende adekvat utredning och behandling har medfört att den norska HTA-organisationen SMM (dvs SBU:s systerorganisation) utarbetat en systematisk översikt över kunskapsläget. Man finner inga stöd för ett kausalt samband mellan typ av skada

och utveckling av specifika symtomkonstellationer.

Objektiva tecken är allt för ospecifika och inte nog robusta för att man ska kunna hävda existensen av en gemensam patofysiologisk mekanism för »kroniskt whiplashsyndrom«.

Rapporten finns att läsa i sin helhet på nätet: www.sintef.no/smm

Ordlista – EBM och klinisk epidemiologi.

II Inom EBM-rörelsen förekommer

många termer från klinisk epidemiologi som kan kännas svårbegripliga och stundtals med andra betydelser än vad man är van vid i sin vardag. Anglicismer är vanliga och deras översättning till svenska kan vara mycket varierande. För att råda bot på detta har docent Jörgen Malmquist för SBU:s räkning sammanställt en ordlista som på ett precist sätt redovisar denna terminologi.

Ordlistan finns fritt tillgänglig och lätt att skriva ut från: www.sbu.se

Träning ger endast blygsam trycksänkning hos kvinnor

II Personer med förhöjt blodtryck får ofta rådet att motionera regelbundet för att uppnå en sänkning av blodtrycket och därmed undvika farmakologisk behandling. Dokumentation från randomiserade studier om långsiktig nytta på överlevnad och komplikationer saknas men en positiv effekt på trycket har antagits trolig.

Fram till 1992 fanns väsentligen bara studier på män men en aktuell metaanalys redovisar de studier som därefter publicerats avseende fysisk träning och viloblodtryck hos kvinnor. En extensiv databassökning fram till 1998 identifierade ca 2 000 studier. I den slutliga analysen inkluderades 10 studier där 732 kvinnor randomiserats till konditionsträning eller en fysiskt inaktiv kontroll-

grupp och där effekten på viloblodtryck uppmätts. Träningsperioden var i snitt ett halvt år med en spridning på mellan 10 och 52 veckor. En fjärdedel av försökspersonerna fullföljde inte träningsperioden. Träningen genomfördes i genomsnitt under cirka 40 minuter, fyra gånger per vecka med måttlig intensitet (40–80 procent av arbetskapacitet).

Deltagarna var väsentligen normotona med initiala tryck på 119/73. Effekten på blodtrycket var blygsam. Det systoliska blodtrycket sjönk med två mm Hg (KI -3; -1) och det diastoliska med en mm Hg (KI -2; -1). Den diastoliska sänkningen var mer uttalad hos yngre kvinnor, men annars kunde inga särdrag urskiljas som skulle prediktera speciellt gynnsam effekt. Såväl vikt som andel kroppsfett och

maximal syrgaskonsumtion förbättrades signifikant i träningsgruppen.

Författaren påpekar den blygsammare effekten hos kvinnor än hos män vad gäller trycksänkning men framhäver andra positiva effekter av träning, oavsett om trycket sjunker eller ej.

De ingående studierna var inte analyserade enligt »intention-to-treat«, uppföljningarna var korta och grupperna små, vilket bidrar till tveksamheter avseende effekten över lång tid och i »verkliga livet«. Största tillkortakommandet är dock att studierna inte utförts på hypertensiva kvinnor där behovet av kunskap är som störst.

Prev Med 1999; 28: 264-75.

Många kirurgiska översikter i senaste Cochranebiblioteket

II I EBM-rörelsens barndom berörde de flesta systematiska översikter behandling med läkemedel, ofta inom internmedicinska ämnen. Så var även fallet inom Cochranesamarbetet vars översikter redovisas i Cochranebiblioteket och nås via www.rondon.se. Bibliotekets databaser uppdateras varje kvartal och helt nya översikter kommer till i varje utgåva.

Årets sista version (Issue 4, 2000) innehåller ca 70 helt nya ämnen varav ett stort antal kirurgiska frågeställningar vilket pekar på att även opererande specialiteter i allt större utsträckning belyses med systematiska översikter och metaanalyser.

Bräckkirurgi berörs i två sammanställningar. Laparoskopiska tekniker jämförs med öppen kirurgi i 34 studier omfattande 6 800 patienter. Slutsatsen är att den laparoskopiska varianten ger mindre postoperativ smärta och snabbare återgång i arbete men tar längre tid att genomföra och möjligen ökar risken för ovanliga men svåra komplikationer [1]. Laparoskopisk kirurgi med användande av nät minskade signifikant recidivfrekvensen (oddskvot 0,56, KI 0,33; 0,93). Öppen kirurgi med eller utan nät jämförs i 15 studier omfattande 4 000 patienter [2]. Komplikationsfrekvensen skilde sig inte mellan grupperna och smärta samt sjukhusvistelse gick inte att utvärdera. Patienter opererade med nät fick mindre ofta recidiv än de som opererades utan nät (oddskvot 0,39, KI 0,25; 0,59).



Cochranebiblioteket kan nås via www.rondon.se

Karpaltunnelsyndrom opereras oftast vid mer permanenta symtom men nyligen har flera randomiserade studier avseende lokal steroidinjektion publicerats. Cochraneöversikten identifierar två relevanta studier med en klar effekt på symtom – jämfört med placebo – vid en uppföljning efter en månad [3]. Längre uppföljningstider och jämförelser med kirurgi eller andra behandlingsformer saknas för närvarande.

Ett försök att klarlägga om kirurgi vid bronkiektasier är till nytta för patienten stupade på den fullständiga avsaknaden av randomiserade eller åtminstone kontrollerade studier [4]. Det är därmed omöjligt att fastslå ingreppets plats vid behandling av patienter med bronkiektasi.

Rädsla för akut nedsättning av njurfunktionen kan bromsa användandet av NSAID-preparat i den postoperativa pe-

rioden där deras analgetiska effekt väl kommer till pass. Det kan då vara lugnande att ta del av resultaten från en systematisk översikt som sammanställt data från åtta randomiserade studier [5]. En kliniskt obetydlig sänkning av kreatininclearance på dag 1 (18 ml/min) normaliserades dag 2 och inga fall av dialyskrävande njursvikt påvisades.

Slutligen jämförs olika uppföljningsstrategier för kvinnor opererade för tidig bröstcancer i en översikt omfattande fyra randomiserade studier och 3 200 kvinnor [6]. Slutsatsen är att uppföljning baserad på enbart klinisk undersökning och årlig mammografi är likvärdig med mer intensiv och avancerad uppföljning mätt såväl som total överlevnad, överlevnad fri från recidiv eller hälsorelaterad livskvalitet.

1. *Laparoscopic techniques versus open techniques for inguinal hernia repair.*
 2. *Open mesh versus non-mesh repair of inguinal hernia.*
 3. *Local corticosteroid injection for carpal tunnel syndrome.*
 4. *Surgery vs non-surgical treatment for bronchiectasis.*
 5. *Effects of nonsteroidal anti-inflammatory drugs on post-operativ renal function in adults.*
 6. *Follow-up strategies for women treated for early breast cancer.*
- (samtliga i Cochrane Library, Issue 4, 2000).

Automatiserade cytologimetoder kan ännu inte ersätta traditionella metoder

II I hela i-världen bedrivs screening i syfte att förebygga cervixcancer. Instrumentet som används, exfoliativ cytologi, har i sin grund varit oförändrat sedan 1940-talet. Metoden är manuell, arbetskrävande, subjektiv och utförs med skiftande kvalitet. En industri byggs nu upp som lanserar komplement och alternativ till traditionell cytologi. Man nöjer sig inte alltid med marknadsföring till professionen. Till exempel i Nya Zeeland går företag ut till allmänheten med uppmaningar att kvinnor ska begära (och själva betala) för prov som påstås erbjuda ett bättre skydd mot cancer. De två mest uppmärksammade och mest marknadsförda komplementen är »liquid-based cytology« (vätskebaseerad cytologi, VBC) och »automated screening« (automatiserad undersökning av cytologprover, AC).

SBU:s systerorganisation i Nya Zeeland, NZHTA, publicerade nyligen en systematisk litteraturgenomgång av metoderna. Översikten gäller arbeten publicerade januari 1997 – maj 2000 och tar vid där ett tidigare arbete, publicerat av Australian Health Technology Advisory Committee, slutade. Med sökstrategin påträffades över 700 artiklar från vilka 26 vaskades fram efter redovisade inklusions- och exklusionskriterier. Tio artiklar värderade ThinPrep och tre den andra VBC-metoden Auto Cyte Prep. I två artiklar utvärderades den enda idag existerande AC-metoden, AutoPap, fem var analyser av kostnadseffektivitet och ytterligare sex översiktsartiklar/metaanalyser. Avsikten var att värdera nyttan av metoderna i Nya Zeelands screeningprogram, som liknar de rekommendationer som Socialstyrelsen utfärdat för de svenska sjukvårdshuvudmännen. Detek-

tionsgränsen sattes till HSIL, vilket med svensk terminologi motsvaras av lesioner med medelsvår dysplasi eller »värrer« (CIN2+).

NZHTA har en hel del metodologisk kritik mot de genomgångna studierna. Ingen av VBC-studierna är randomiserad. De flesta har använt historiska kontroller eller kontroller i andra populationer med okänd eller tveksam jämförbarhet, vilket kan ge en rad olika bias. Två studier har försökt komma förbi detta genom att låta ta två prov från samma kvinna, något som dels missgynnar det sist tagna provet, dels har proven tagits ur högriskpopulationer vilket bl a inte tillåter beräkningar av prediktiva värden som är relevanta för populationsbaserad screening. Industrifinansiering har ett negativt meritvärde.

En genomgående kritik av alla originalartiklar utom en är att de saknar verifiering av negativa prov. Detta omöjliggör exakta analyser av specificitet och sensitivitet. Författaren medger samtidigt de etiska och praktiska problem som kolposkopi och biopsier på kvinnor utan cytologatypier innebär. Expertpaneler med oberoende cytologer, som granskar negativa prov, är den något tveksamma nödlösning som NZHTA för fram. Histologi är som alltid i dessa sammanhang golden standard men dessvärre undviks sorgfälligt det metodologiska problemet att biopsier är beroende av en undersökningsmetod (kolposkopi) med sensitivitet och specificitet som är jämförbar med just cytologi.

Automatiserad cytologi (AutoPap) hade undersökts kombinerat för kvalitetskontroll och primärundersökning. I det senare fallet fick maskinen sortera ut de mest normala proven, maximalt 25

procent behöver då inte undersökas manuellt. Det förelåg ingen skillnad i detektion av HSIL+.

Arbetet har en förtjänstfull diskussion kring kostnadseffektivitet och tar in begreppet »societal cost« där även livskvalitetsaspekter hos kvinnor som deltar i screeningprogrammet ingår, exempelvis tidsspillan och oro vid besked om avvikande provsvar.

Konklusionen är att några slutsatser om metodernas sensitivitet och specificitet inte kan dras, och att det sålunda inte finns några bevis för förbättrad detektion av höggradiga (CIN2+) förändringar. Kostnadseffektiviteten, som är beroende av metodernas sensitivitet och specificitet, blir därmed också svårbedömd. Ju sämre laboratoriestandarden är, desto större är sannolikheten att man kan påvisa en förbättring med de nya metoderna. Introduktion av nya screeningmetoder i avsikt att förbättra sensitivitet riskerar att göra detta på specificitetens bekostnad – screeningens Skylla och Karybdis. Små försämringar av specificitet kan ge dramatiska fördyringar. Traditionell cytologi, som enligt rapporten förebygger 93 procent av cervixcancerfallen vid screening vart tredje år och fullt deltagande, rekommenderas vara basen för screening mot denna sjukdom även i fortsättningen.

Björn Strander

bjorn.strander@swipnet.se

Broadstock M. Effectiveness and cost effectiveness of automated and semi-automated cervical screening devices – A systematic review of the literature. NZHTA Reports, 2000; Issue 3 (<http://nzhta.chmeds.ac.nz/csv3n1.htm>)

I korthet

Evidensbaserad medicin har försummat biverkningarna

II De systematiska översikter som utgör ryggraden i EBM handlar oftast om att utvärdera behandlingar med fokus på kliniskt robusta effektmått såsom överlevnad, komplikationer eller kostnadseffektivitet. Kunskap om behandlingarnas säkerhet är lika viktiga när evidensbaserade beslut ska fattas avseende värdet av en viss metod, såsom läkemedel, kirurgi, psykologiska eller fysikaliska behandlingar. Ändå finner läsaren biverkningar och andra baksidor av behandlingar styvmoderligt behandlade i

tunga kunskapsunderlag, såväl Cochranöversikter som SBU-rapporter. En genomgång av systematiska översikter noterade i Medline styrker denna uppfattning.

Under femårsperioden 1986–1990 redovisades och värderades »harm/safety« i 3 procent av översikterna. Under nästa femårsperiod ökade andelen till 6 procent för att de senaste åren hamna på ca 1 procent.

Även om det skett en viss förbättring är läget fortfarande mycket otillfredsställande. Författarna pekar ut problemet som sammanhänger med att de systematiska översikterna strikt begränsas till randomiserade kliniska undersök-

ningar. Det är dock välkänt att en stor del av problem med behandlingsmetoder inte kommer att identifieras under korta behandlingstider i begränsade grupper. Lösningen är att för bedömning av »harm« inkludera data från t ex fallrapporter, farmakovigilansstudier och »post-marketing surveillance«. Utan resultat från sådana komplementära studiedesigner kan aldrig den kritiska frågan besvaras på ett rigoröst sätt: gör denna behandlingsmetod mer nytta än skada?

Ernst E, Pittler MH. Systematic reviews neglect safety issues. Arch Intern Med 2001; 161: 125-6

Ge låg dos salicylika till gravida kvinnor med ökad risk att få preeklampsi!

II Preeklampsi (blodtrycksstegring och proteinuri) är en fruktad komplikation under graviditet, som medför stora hälsorisker för såväl mor som barn. Om tillförsel av låga doser salicylsyra skyddar mot utvecklingen av preeklampsi hos kvinnor med ökad risk att få sjukdomen har varit oklart. Flera mindre studier visade tidigt tydliga positiva effekter, men i den största randomiserade studien (CLASP) kunde någon säker effekt inte konstateras. I en aktuell systematisk översikt har alla randomiserade studier där låg dos salicylika jämförts med placebo eller ingen behandling identifierats. Studiernas validitet bedömdes enligt riktlinjer från Cochrans handbok. Trehundratio rapporter identifierades. Efter bedömningen återstod 39 studier, omfattande 30 563 kvinnor, som användes för metaanalysen.

Risken att få preeklampsi minskade med 15 procent (relativ risk 0,85; konfidensintervall 0,78–0,92) hos behandlade kvinnor. Etthundra kvinnor behövde behandlas för att förhindra ett fall av preeklampsi. Användes högre dos salicyli-

ka, mer än 75 mg, minskade risken än mer (relativ risk 0,35), men endast 1 264 kvinnor ingick i dessa försök. Risken att föda för tidigt (före 37 veckor) reducerades med 8 procent. Sjuttio två kvinnor måste behandlas för att en förtidsbörd skulle förhindras. Mortaliteten för barnen minskade med 14 procent. I de studier där behandlingen påbörjats före 20:e graviditetsveckan respektive mer än 75 mg salicylsyra getts, noterades en signifikant minskad förekomst av SGA-barn (små för tiden). För övriga undersökta utfall, bl a förekomst av blödningar hos barnen, sågs inga skillnader.

Författarna konkluderar att hos kvinnor med förhöjd risk att få preeklampsi medför låg dos salicylika att risken minskar med 15 procent. Dessutom innebär behandlingen en reduktion av dödföddhet, neonatal död samt förtidsbörd. Möjligen kan doser på mer än 75 mg medföra än större positiva effekter. För sådana doser är dock kunskapen om säkerhetsaspekterna inte lika solid. Det finns inga säkra data, som besvarar frå-

gan, när i graviditeten behandlingen bör startas. Författarna ämnar fortsätta sitt arbete för att söka kunskap om vilka kvinnor, som kan tänkas ha störst nytta av profylaktisk behandling med låg dos salicylika under graviditet.

För närvarande gäller rekommendationen att kvinnor med ökad risk att utveckla preeklampsi bör behandlas med låg dos (<75 mg) salicylika. Även om vinsterna är måttliga är de värdefulla ur ett folkhälsoperspektiv. Att påbörja behandlingen före 12 veckor eller att använda högre dos än 75 mg kan inte rekommenderas innan ytterligare kunskap vunnits beträffande säkerhetsaspekterna.

Ove Axelsson

ove.axelsson@kbh.uu.se

Duley L et al. Antiplatelet drugs for prevention of pre-eclampsia and its consequences; systematic review. BMJ 2001; 322: 329-33 (<http://www.bmj.com/cgi/content/abstract/322/7282/329>)

Minskad postoperativ mortalitet och morbiditet vid operation i regionalanestesi

II En systematisk översikt med metaanalys av effekten av regionalanestesi (RA) på postoperativ mortalitet och morbiditet i samband med kirurgi publicerades i BMJ i december.

Den baseras på 141 publicerade studier med 9 559 patienter. Studierna är utvalda på basis av adekvat randomisering och att fullständiga data avseende 30 dagars-mortalitet och -morbiditet kunde erhållas från respektive provare. Litteraturtäckningen är i det närmaste total och studiedesignen likartad i de olika studierna. Validiteten av data är hög, noga kontrollerad via centra i varje land, där författarna vid behov fått tillgång till originaldata från respektive provare. Man har delat in materialet i RA/icke-RA, oavsett om patienterna dessutom erhållit generell anestesi (GA). När till-

räckligt underlag funnits har subanalys gällande operations- och anestesityp utförts.

Trettio dagars mortalitet var 2,1 procent efter RA och 3,1 procent vid icke-RA, dvs en relativ riskminskning med 1/3, $P=0,006$, eller en absolut riskreduktion på 1 procentenhet. Man fann ingen skillnad mellan olika typer av kirurgi, användning av spinal eller epiduralanestesi, eller om RA kombinerades med GA. Däremot var mortaliteten lägre om RA användes även postoperativt. Allvarligare komplikationer såsom djup ventrombos, lungemboli, infektioner, hjärtinfarkt, njursvikt, andningsdepression och blödning var signifikant lägre vid RA än GA. Man kan dock inte dra slutsatser om orsaken till det positiva utfallet beror på användningen av RA eller

om det beror på undvikande av komplikationer till GA.

Användning av RA är vanlig som standard inom elektiv kirurgi i Sverige, men fördelarna i postoperativ vård kan sannolikt utnyttjas i högre grad än nu. En diskussion om risk-nyttaförhållandet pågår ständigt och under senare tid har riskerna betonats påtagligt i den inhemska debatten.

Märta Segerdahl

marta.segerdahl@anaesth.hs.sll.se

Rodgers A et al. Reduction of postoperative mortality and morbidity with epidural or spinal anaesthesia: results from overview of randomised trials. BMJ 2000; 321: 1493-7 (<http://www.bmj.com/cgi/content/abstract/321/7275/1493>)

Skriv till oss!

Läkartidningen vill uppmuntra läsarna att skriva kommentarer och referat av systematiska översikter med relevans för en bredare läsekrets. Uppslaget »Nya rön« tar upp evidensbaserad medicin med betoning på referat av bra systematiska översikter och metaanalyser. Andra aktuella artiklar som berör evidensbaserad medicin kan också vara av intresse för läsarna.

Om du vill skriva kontakta Mats Eliasson, medicinsk redaktör. Mats.Eliasson@nll.se eller 070-513 02 93.

Mottagningen tar längre tid med datoriserade journalsystem

■ Ett deskriptivt överflöd, men grav brist på utvärdering, är betyget engelska forskare avger efter att systematiskt ha sammanställt studier avseende datoriserade journalsystem i primärvården [1]. Endast ett fåtal studier var randomiserade, flera saknade kontrollgrupper eller var utförda med en före-efterdesign. Uppenbart var att tiden för mottagningsbesöket förlängdes och ökade läkarcen-trerad verksamhet på bekostnad av patientinriktad aktivitet.

Datorsystemets påminnefunktioner, s k reminders, ökade däremot genomförande av vaccinationer och screening för cervixcancer och högt

blodtryck. När elektroniska protokoll användes vid kontroll av patienter med hypertoni eller diabetes fann man påtagliga förbättringar i själva processen men återigen till priset av längre tid. Elektronisk recepthantering förbättrade förskrivning av generika och ledde till lägre kostnader. Endast 17 studier mätte någon patientrelaterad effekt, och fann därvid bl a lägre systoliskt tryck hos hypertoniker, men resultatet från antikoagulationsmottagningar var motsägelsefulla. Remisser till specialistvård ökade inte och inte heller förändrades patienttillfredsställelsen vid användande av datoriserade system.

Sammanfattningsvis finns vissa indikationer på att datoriserade journalsystem med påminnefunktioner och beslutsstöd för förskrivning av recept samt provbeställning kan förbättra vården men tar längre tid. Patientnyttan är tills vidare dåligt bevisad.

Mitchell E, Sullivan F. A descriptive feast but an evaluative famine: systematic review of published articles on primary care computing during 1980-97. BMJ 2001;322:279-82 (<http://www.bmj.com/cgi/content/full/322/7281/279>)

AT1-blockerare – upp till evidens!

■ Kostnaderna för AT1-blockerare (ATB) ökade mellan år 1999 och 2000 med drygt 40 procent, från 176 till 251 miljoner kr/år. Försäljningen, mätt som DDD/1 000 invånare, ökade fyra gånger mellan 1996 och 1999 medan ACE-hämmarna väsentligen låg kvar på samma nivå. Förskrivningen sker huvudsakligen för hypertoni, och i viss mån också för hjärtsvikt. Högst är förskrivningen i Uppsala län, tätt följt av Gävleborg, Dalarna, Norrbotten och Västmanland och i särklass lägst på Gotland. Men hur är det med dokumentationen av ATBs kliniska effekt?

Norman Kaplan vid University of Texas föreslår i en kritisk kommentar att man bör hålla marknadsförarna stången med frågan: »Show me the outcome data.« Han konstaterar att endast en minoritet av de 23 miljoner amerikaner som behandlas med preparaten gör det på grund av intolerans för ACE-hämmare, den indikation som alla tunga riktlinjer för hypertoni-behandling anger.

Anledningen till hans skepsis är till att börja med den totala bristen som föreligger avseende »hårda end-points«, dvs data avseende effekt på död, stroke eller hjärtinfarkt, från randomiserade kontrollerade studier med dessa läkemedel. Som ett varnande exempel på faran med att lägga alltför stor vikt vid surrogatmått tar han alfablockerare som i många fall saluförts som ett speciellt bra medel vid hypertoni till följd av gynnsamma effekter på sådant som insulinkänslighet, fibrinolys och lipider.

Men i den första stora randomiserade studien där doxazosin jämfördes med bl a ett enkelt tiaziddiuretikum (ALLHAT) tvingades man bryta studien då doxazosinpatienterna utvecklade hjärtsvikt dubbelt så ofta som diuretikapatienterna. Kaplan m fl är mycket kritiska till att amerikanska FDA godkänner hypertoni-preparat som t ex ATBare enbart på basen av surrogatdata. Farorna med att lita på surrogatmått belyses vidare i en kommande artikel i EBM-serien.

Den andra anledningen till en mer avvaktande attityd till ATB är studien som jämförde losartan med kaptopril vid hjärtsvikt (ELITE II). Studien visade en icke-signifikant lägre mortalitet hos patienter som behandlats med ACE-hämmare jämfört med dem som behandlats med ATBare. Studien som ibland anförts som ett bevis för att preparaten är likvärdiga kan inte tas som intäkt för ekvivalens. Sådana studier kräver betydligt större patientmaterial och fördefinierade minsta viktiga skillnad i effektmåten (se Läkartidningen 2000, nr 32-33, sidan 3453).

I väntan på resultat från ett stort antal pågående studier med relevanta effektmått bör alltså ATB-preparat enbart användas på mycket speciella indikationer, vilket knappast är fallet i Sverige.

Kaplan N. Should new drugs be used without outcome data? Implications of ALLHAT and ELITE II. Arch Intern Med 2001;161:511-2

Visst stöd för vitlökenes lipidsänkande effekt

■ Den ursprungliga skaparen av den systematiska översikten som en vetenskaplig metod, Cynthia Mulrow, står bakom en aktuell kritisk sammanställning av 45 randomiserade kontrollerade studier där vitlökenes kardiovaskulära effekter studerats. Studierna var ofta av kort varaktighet, tveksamt randomiserade, dåligt placebokontrollerade och med stor andel av patienterna som avbröt. Upp till tre, men inte sex månader, fanns en viss sänkning av kolesterol, LDL och triglycerider. HDL-nivåer påverkades inte. Effekten på blodtrycket var motsägelsefull men däremot fanns

ingen inverkan på glukos- eller insulinbalansen. Viss minskning av trombocyttaggregation påvisades. Ingen studie med adekvat design hade studerat effekter på mortalitet eller kardiovaskulär sjuklighet. Problem med obehaglig an-

dedräkt eller doft var helt klarlagda och dessutom förekom en hel del gastrointestinala klagomål.

Det finns således ett svagt stöd för små och gynnsamma effekter i ett kort perspektiv, men forskningen kring vitlök har många hinder att övervinna innan dess eventuella plats i den terapeutiska arsenalen är klarlagd.

Ackerman R et al. Garlic shows promise for improving some cardiovascular risk factors. Arch Intern Med 2001;161:813-24. (<http://archinte.ama-assn.org/issues/v161n6/abs/ira00067.html>)

Oklart om egen blodsockermätning förbättrar metabol kontroll hos patienter med typ 2-diabetes

II Egenmätning av blodsocker framhålls ofta som en hörnsten vid behandling av diabetes, både typ 1 och typ 2. Förhoppningen är att sådan mätning ska ge patienterna återkoppling och leda till justering av farmakologisk behandling, kost och motion samt minska riskerna för farliga hypoglykemier. Detta bör då avspeglas i förbättring av den glukosmetabola kontrollen, mätt som HbA1c och även i ökad livskvalitet för att slutligen omsättas i lägre risk för mikro- och makrovaskulära komplikationer samt minskad mortalitet.

Under år 2000 förskrevs testremisor för blodglukosmätning till ett pris av 423 miljoner kr i Sverige varav hela 196 miljoner kr berörde patienter 65 år eller äldre, dvs till allra största delen patienter med typ 2-diabetes. Urinstickor för glukostestning har nästan försvunnit helt och såldes för endast drygt en miljon kronor. Men uppnår vi verkligen målsättningen med egentestningen eller kunde patientens och sjukvårdens insatser användas bättre?

Engelska forskare redovisade nyligen en systematisk sammanställning av de åtta studier som randomiserat typ 2-diabetiker till egentest av blod- eller urinsocker eller till ingen testning alls [1]. Patienterna var i 60-årsåldern och behandlades med kost eller peroral preparat, endast en studie inkluderade insulinbehandlade. Studierna var alltför små för att ha statistisk styrka för att påvisa skillnader i HbA1c mindre än 0,5 procent och uppföljningstiderna var korta

med stora bortfall. Endast en av studierna påvisade en signifikant, men blygsam, reduktion av HbA1c och vid en metaanalytisk sammanvägning av fyra studier (285 patienter) var glykerat hemoglobin 0,25 procent lägre (95 percents konfidensintervall -0,6 till 0,1) i gruppen med regelbunden egenmätning jämfört med dem som inte mätte socker. Tre studier med 278 patienter jämförde mätning av sockernivåerna i urin med mätning i blod men fann ingen skillnad i glykerat hemoglobin. Man fann inga skillnader i förekomst av hypoglykemi, livskvalitet, välmående eller vikt mellan dem som mätte socker och dem som inte alls testade.

De bakomliggande studierna var generellt av dålig kvalitet och resultaten har breda konfidensintervall, varför man inte kan utesluta vare sig en gynnsam förbättring eller en skadlig försämring av HbA1c-nivåerna. Resultaten är dock i överensstämmelse med två tidigare översikter i frågan. Det bör betonas att utvärdering av enskilda komponenter inom så komplexa interventioner som diabetesvård är mycket svårt och att forskningsmetoderna behöver utvecklas. Ändå drar författarna slutsatsen att det inte finns anledning att uppmantra patienterna att köpa dyra testapparater eller utnyttja viktig mottagningstid till att lära patienterna testa och analysera deras resultat. Den sjukvårdsekonomiska aspekten är ju inte heller försumbar.

Man kan bara instämma i deras krav på en stor randomiserad studie med lång

varaktighet för att utröna effekterna av egentest, inte bara på blodsockerbalans utan även på komplikationer, hälsorelaterad livskvalitet och hälsoekonomi. En sådan studie bör ha en väl standardiserad utbildnings- och träningskomponent och inkludera en majoritet insulinbehandlade patienter med typ 2-diabetes. Är det någon av testremisefabrikanterna som är beredd att initiera denna studie?

I detta sammanhang vill jag peka på ett mycket välskrivet och koncist debattinlägg till stöd för en vård av typ 2-diabetiker som verkligen vilar på en solid vetenskaplig grund, dvs huvudsakligen randomiserade och kontrollerade studier [2]. Richter och Berger från WHO's Collaborating Centre for Diabetes och Cochranes diabetesgrupp pekar på en mångfald obesvarade frågor, t ex nyttan av de nya preparaten som speciellt ska påverka postprandiella blodsockerstegringar, ett hitintills helt obevisat surrogatmått.

1. Coster S et al. Self-monitoring in type 2 diabetes mellitus: a meta-analysis. *Diabet Med* 2000;17:755-61
2. Richter B, Berger M. Randomized controlled trials remain fundamental to clinical decision making in type II diabetes mellitus. *Diabetologia* 2000;43:528-32

Till slut ...

II För ett år sedan startade Läkartidningens satsning på evidensbaserad medicin. För detta ändamål avsatte vi var fjärde vecka uppslaget Nya Rön till korta presentationer av bra och kliniskt relevanta systematiska översikter och viktiga EBM-artiklar.

Avsikten var att visa vilken nytta man som vanlig kliniker har av att tänka och agera evidensbaserat (glöm inte att ställa frågan »Show me the evidence«!), samt visa på den rikedom av valid medicinsk kunskap som ryms i systematiska översikter; från Cochrane, HTA-organisationer och tidskrifter. Efter 11 uppslag och presentation av drygt 50 översikter inom vitt skilda områden avslutar vi denna specialutgåva av Nya Rön.

Jag tackar mina medförfattare och sammanfattar det ofta nedslående budskapet att den nuvarande vetenskapliga basen för klinisk medicin ofta är bräcklig och sällan ger stöd för så tvingande rekommendationer som vi möter i vårdprogram och riktlinjer. Lösningen är inte i första hand fler metaanalyser utan att klinisk forskning sker med bättre metoder, större patientgrupper, längre uppföljningar och att effekten mäts med mått som är relevanta för patienten. Det är centralt med forskningsfrågeställningar som inte styrs av industrins lanseringsbehov men som utgår från ett patientperspektiv.

EBM-serien fortsätter dock året ut med

artiklar om surrogatmått, klass-effekter och hur man kritiskt läser artiklar om livskvalitet, hälsoekonomi, diagnostik och metaanalyser. Vi hoppas också kunna presentera etiska och idéhistoriska aspekter på evidensbaserad medicin. Självklart välkomnar vi fortfarande egna referat av bra EBM-litteratur men även artiklar som kan problematisera evidensbaserad medicin från egna kliniska erfarenheter och funderingar över vad som kommer efter evidensbaserad medicin.

Mats Eliasson
medicinsk redaktör
evidensbaserad medicin

AVSLUTNING



ebm

LÄKARTIDNINGENS SERIE OM EVIDENSBASERAD MEDICIN 2000-2002

Lars Werkö, professor, tidigare styrelseordförande, Statens beredning för medicinsk utvärdering (SBU), Stockholm

Kjell Asplund, professor, Umeå; styrelseordförande, SBU

Peter Aspelin, professor, Huddinge Universitetssjukhus; ordförande, SBU:s råd

Mona Britton, professor, medicinskt sakkunnig, SBU

Mats Eliasson, överläkare, Luleå; projektsamordnare, SBU; medicinsk redaktör, Läkartidningen (evidensbaserad medicin)

Jean-Luc af Geijerstam, doktorand, SBU-medarbetare

Sten Thelander, sakkunnig i psykiatri, SBU

Två år med EBM i Läkartidningen

Klinisk forskning och rutinsjukvård har närmat sig varandra

|| Vi har nu levt med begreppet evidensbaserad medicin, EBM, i några år. Några läsare har varit entusiastiska, några tveksamma eller negativa, medan flertalet har varit avvaktande.

Det kan finnas anledning att försöka bedöma vad EBM står för, vad som hittills har åstadkommit och vilka svagheter eller vilken styrka som begreppet innehåller. Framför allt är det viktigt att se hur begreppet kan utvecklas, vad som saknas och vilka nya vägar som måste beträdas för att nå målet: bästa möjliga sjukvård med optimalt använda resurser. Det gäller att vara lika kritisk till detta begrepp som till alla andra försök att snöra in den kliniska medicinen i en tvångströja. Begreppet måste utsättas för en ständig kritisk diskussion och omprövning.

Är EBM något nytt?

EBM är egentligen endast en omformulering av devisen »vetenskap och beprövad erfarenhet«. Den medicinska utvecklingen går snabbt och nya möjligheter erbjuds så gott som varje dag för både diagnos och behandling. När dessutom aktuell medicinsk kunskap i allt större utsträckning finns tillgänglig på elektroniska medier behövs det en annorlunda definition av vad som är vetenskap i det kliniska sammanhanget.

Benämningen vetenskap har därmed blivit bättre definierad och innebär i dagens vokabulär att aktuell vetenskaplig litteratur skall tjäna som underlag för det medicinska handlandet. Man måste således systematiskt söka i tillgängliga vetenskapliga databaser för att vara säker på att man har rätt uppfattat den senaste vetenskapliga ståndpunkten. I många fall kan det räcka att konsultera systematiska litteraturoversikter utarbetade på ett tydligt sätt med sökstrategier, kritisk granskning, kvalitetsbedömning och tydlig syntes. Mest tilltro kan man i allmänhet ha till obundna gruppers arbete och

SAMMANFATTAT

EBM, evidensbaserad medicin, har

- återupprättat den kliniska patientnära forskningen; inga teoretiska eller experimentella studier kan ersätta kliniska observationer med utfall som är av direkt betydelse för patienten.
- återupprättat den oberoende vetenskapliga granskningen, fri från allmänt tyckande, prestige och kommersiella intressen.

Evidensbaserad medicin

Se även artikeln på sidan 3492 i detta nummer.

till myndigheter där olika experter och synsätt varit representerade.

Hur är det med erfarenheten?

Även det andra ledet i läkardevisen, det som betonar erfarenheten, har fortfarande betydelse. Det är ett missförstånd att tro att EBM inte längre skulle röra det vi kallat »beprövad erfarenhet«. Men vad är beprövad erfarenhet? Den vane klinikern, som sett många exempel på hur en sjukdom kan te sig hos olika patienter, besitter ett underlag för hur ett visst tillstånd kan utvecklas och hur det kan påverkas. Denna erfarenhet har byggts upp genom aktivt lyssnande och frågande, genom analys, undersökningsfärdigheter, omdöme och förmåga till syntes. Allt kan förbättras med ökad erfarenhet, som

många gånger har utgjort ett väsentligt bidrag till det beslut som fattats i olika situationer. Erfarenheten är dock sällan systematisk; den kan färgas både av det spektakulära och av det som hänt senast.

Den rätta användningen av den egna erfarenheten borde vara att i samband med analys av ett aktuellt problem ställa erfarenheten mot vad litteraturen visar. Sammanvägningen av den egna bilden med litteraturens kan sedan presenteras för patienten, som också måste få ta ställning till de åtgärder som föreslås. Detta kan låta krångligt, men innebär endast ett betydligt merarbete – när man väl vant sig att tänka i dessa banor.

Vilka fordringar ställer EBM på klinisk forskning?

Idén med EBM innebär ett närmande inte bara mellan den kliniska patientnära forskningen och rutinsjukvården, utan också mellan den teoretiska biomedicinska forskningen och den patientnära forskningen. Det är ju först när teoretiska fynd eller teorier fått vetenskaplig tillämpning på patienter som resultat av värde för sjukvården kan åstadkommas. Detta kommer inte att ske om inte de som är ansvariga för forskningen inser vad som fordras i fråga om tillämpad forskning. Mycket av vad som publiceras tycks ha tillkommit mer för forskningens egen skull än för att bilda underlag för en praktisk medicinsk verksamhet.

Erfarenheten har visat att en mycket stor del av den litteratur som finns i databaserna inte duger för att besvara praktiska spörsmål, som t ex hur man skall ta hand om en patient i ett visst läge. Kliniska undersökningar har hittills i alltför stor omfattning givit efter för kravet på statistisk bearbetning, vilket har inneburit att patientmaterialet har renodlats. Det behövs fler studier av patienter som motsvarar dem vi ser i rutinsjukvården, dvs patienter i högre åldrar eller med fler komplicerande sjukdomar än vad som i allmänhet är fallet i de studier som ofta läggs upp på ett sätt som skall underlätta registrering (eller marknadsföring) av ett nytt läkemedel.

Det är därför viktigt att forskningsledare och anslagsbeviljare inser behovet av bra underlag för beslut i rutinsjukvården. Det gäller också att studera vanliga fenomen i praktisk sjukvård i stället för att mäta det som är lätt att mäta, dvs surrogat-händelser och laboratorievården. Att finna vägar för att göra vetenskap av kvalitativa bedömningar kommer också att få större betydelse.

Det är också viktigt att bredda kravet på kontrollerade undersökningar till andra områden är de som varit av störst intresse hittills. Ett sådant är den patologiskt anatomiska diagnosen, som skulle kunna förbättras genom att i större utsträckning tillämpa evidensbaserad patologi [1].

Stimulering genom utbildning och kvalitetsuppföljning

Om EBM skall kunna utvecklas måste vården organiseras på ett sätt som ger möjlighet för läkare och annan vårdpersonal att följa den aktuella vetenskapliga utvecklingen. Vi menar då inte deltagande i kongresser eller liknande aktiviteter, utan att den dagliga verksamheten måste planläggas så att den medger tid till utbyte av kunskap och erfarenhet och till reflektion. EBM innebär ett krav på kontinuerliga utbildningsinsatser.

Det ökande antalet kvalitetsregister innebär en möjlighet att kontinuerligt följa upp verksamheten. Här har ortopedin varit något av ett föredöme; man har nu möjlighet att jämföra operationstekniker och kliniker över hela landet.

Också inom andra discipliner ger kvalitetsregistren möjlighet att göra jämförelser. Det gäller att redovisa hur man utnyttjar metoder som är evidensbaserade, med syfte att förbättra och ständigt följa verksamheten. Inom internmedicinen kan registreringen avse andel patienter i olika åldrar som får trombolys, andel slaganfallspatienter som vårdas på special-

enhet, frekvens uppföljda biverkningar vid antikoagulansbehandling eller hur patienter med diabetes kontrolleras.

De grupper för kollegialt utbyte som allmänläkarna har organiserat i form av FQ-grupper, sjukhusklinikernas »journalklubbar« eller regelbunden diskussion av fall är andra exempel på aktiviteter som behöver stimuleras om den nya kunskap som den patientnära forskningen åstadkommer skall kunna spridas. Om sådana aktiviteter dessutom kan omfatta flera professionella grupper är chansen till ökad kvalitet i vården desto större. Det är sannolikt att särskilda insatser behöver göras för att lära ut hanteringen av den nya informationsteknologin [2].

Här har både huvudmän och kollegiala organisationer ett stort ansvar. Vi befinner oss fortfarande i början av en period där EBM kommer att få allt större genomslag. Den kommande utvecklingen blir beroende av både hur de »vanliga« sjukvårdsarbetarna kan ta upp nya idéer och hur sjukvårdens huvudmän reagerar på kravet på mer tid för reflektion i den dagliga rutinen. Vidare är EBM en nödvändig grund för all prioritering; minst kontroversiell, förstås, när det är fråga om att slopa det som inte visats göra någon nytta.

Implementering i praxis

Ett sätt att utnyttja evidens i klinisk praxis är att arbeta med evidensbaserade vårdprogram, som uppdateras varje till vart tredje år beroende på utvecklingen. Dessa vårdprogram torde kunna täcka högst 80 procent av verksamheten. För övriga 20 procent får man leta i litteraturen efter fakta. Detta kan ofta göras i samband med den rutinmässiga undervisningen (av kandidater, AT- eller ST-läkare) vid kliniken. Om man inte vet och inte finner något i litteraturen får man diskutera den vidare strategin. Det kan då vara viktigare med exspektans, »watchful waiting«, och empati än att till varje pris vara aktiv, med risk för att åstadkomma mer skada än nytta. Särskilt problematiskt kan det bli när man inte når någon säker diagnos (vid yrsel utan säker orsak, oklar bröstsmärta, trötthet). Även om den vetenskapliga litteraturen i sådana fall är sparsam kan den dock ge vissa förslag till optimal handläggning.

För den vane klinikern är det en omdömesfråga om hur ofta man behöver gå till en databas för att konfirmera vad man redan vet (eller snarare tror sig veta). För många kan det bära emot att ständigt ifrågasätta (sitt eget kunnande), för andra kan det bli en helt naturlig sak att vara öppen mot världen. Några kan hävda att den tidspress som dagens kliniker lever under inte medger denna ständiga kunskapskontroll. Så kan ofta vara fallet i dagens rutin, men då är arbetet inte organiserat så att det möjliggör högsta möjliga kvalitet. Och det är vad patienten har all anledning att fordra.

Den rörelse för EBM som nu har etablerats sedan något tiotal år tillbaka har producerat mycket litteratur om begreppet EBM: böcker, artikelserier i internationella tidskrifter och koncentrerade anvisningar om vad EBM är, hur man skall söka litteratur och överhuvudtaget tillämpa idéerna i det dagliga arbetet. Antalet systematiska översikter, metaanalyser och riktlinjer har ökat nästan exponentiellt. Det finns således stora möjligheter och alltmer lätthanterliga metoder för den intresserade läkaren att använda sig av EBM.

Naturligt med ett motstånd mot kontroll

Ingen tycker att det är trevligt att ha en kritiker stående bakom ryggen kontrollerande vad man gör, i synnerhet inte om man är en etablerad specialist som själv kan bedöma den verksamhet man sedan länge bedriver. Det finns därför ett naturligt motstånd mot att ständigt bli utvärderad (eller kontrollerad), vare sig det kallas klinikrevision, audit eller teknologitvårdning. Det gäller särskilt om den som skall utvärdera kommer från en annan yrkesgrupp eller en annan kultur än

den man själv tillhör, eller – kanske ännu mer motbjudande – från en myndighet.

Kontinuerlig kontroll och utvärdering kräver dock att inte bara de egna, de närmaste kollegerna, skall utvärdera verksamheten. Det fordras också att andra med annan bakgrund, annan utblick och skepsis deltar i bedömningen. Det fordras således en ny form (eller ökad grad) av öppenhet mellan olika specialiteter, mellan specialister av olika slag, för att uppnå en rutinsjukvård baserad på solida kunskaper i stället för på önsketänkande. Denna öppenhet måste då kombineras med en självdisciplin som varken specialister eller andra sjukvårdsarbetare varit särskilt vana vid eller angelägna om. Den är emellertid nödvändig om den kliniska medicinen skall kunna ta del av de framsteg som nu snabbt görs i många naturvetenskapliga eller biomedicinska laboratorier.

Bäst gagnas en positiv utveckling av internt utvecklingsarbete med vårdprogram, resultatuppföljning, omvärderingar och kontinuerlig fortbildning. En sådan styrning av verksamheten upplevs mest utvecklande för alla inblandade.

Det bästa får inte bli det godas fiende

Det är dock en fysisk omöjlighet att i allt handlande tillämpa den rutin som skisserats ovan. Det innebär att EBM egentligen mer är en fråga om en filosofi än detaljerade föreskrifter för hur den dagliga sjukvården skall bedrivas. Ett kritiskt förhållningssätt till gamla och nya metoder, till uttalanden av experter, lärare och professorer är den viktigaste ingrediensen i denna filosofi. Fråga alltid »Var står det?« och kontrollera att det finns ett underlag för de svepande påståenden som auktoriteter alltid har haft en tendens att strö omkring sig.

Utvecklingen kommer att vara beroende dels av de attityder som sprids inom sjukvården, dels av de möjligheter som informationsteknologin kommer att ställa till förfogande. Redan den grundläggande utbildningen måste förmedla den rätta inställningen till hur kunskap skall inhämtas, behållas och utvecklas. Den måste uppmuntra ett kritiskt tänkande och betona att vi inte kan veta allt, men att mycket finns tillgängligt i databaser.

Risker

Det finns dock en risk för att utvecklingen kan komma att drivas och behärskas av kommersiella krafter. De senaste årens ökande beroende av företagsstöd när det gäller vidareutbildning och klinisk forskning är en varningssignal. De medicinska fakulteterna och olika specialistorganisationer måste bevaka att det blir de medicinskt vetenskapliga intressena som fortsätter att leda utvecklingen.

Det gäller också att se till att de olika organ för utvärdering av medicinska metoder som nu är verksamma runt om i världen kan fortsätta sitt arbete med stöd från både huvudmän och sjukvårdspersonal.

Att professionen själv måste ta ansvar för framtiden utan att bli beroende av vare sig sin egen prestige eller kommersialism kan tyckas självklart. Kan universiteten, forskarna och överheten motstå frestelsen att låta »marknaden« styra, med de nackdelar man kan förutse i detta sammanhang?

*

Potentiella bindningar eller jävsförhållanden: Inga uppgivna.

Referenser

1. Fleming KA. Evidence based cellular pathology. *Lancet* 2002;359: 1149-50.
2. Regeringskansliet. Vård i tiden. Strategier och åtgärder för att bredda användningen av telemedicin och distansöverbyggande vård. Stockholm: Socialdepartementet; 2002. DS 2002:3.

EBM-serien avslutas

■ Läkartidningens artikelserie om evidensbaserad medicin (EBM) har sedan mars år 2000 publicerats som återkommande temanummer. Den allra första artikeln »EBM – ett konkret redskap för vårdens verkstadsgolv« av Lars Werkö speglade seriens målsättning att påvisa nyttan av ett evidensbaserat synsätt som ett viktigt hjälpmedel för att fatta väl underbyggda kliniska beslut.

I serien har vi publicerat 32 metodartiklar om EBM, 28 kommentarer till systematiska översikter, flera debattinlägg och ett stort antal evidensbaserade Nya Rön. Enligt en nyligen genomförd läsundersökning har EBM-artiklarna varit bland de mest lästa och uppskattade delarna av tidningen, inte minst bland läkare under utbildning.

Att Läkartidningen satsat på att främja ett vetenskapligt förhållningssätt till klinisk medicin har således fått positiva läsreaktioner, och intresset för serien visar att det finns ett behov av att tillägna sig en metodik som gör det möjligt att bedöma om underlaget för klinisk praxis bygger på väldefinierade studier med adekvat design och där relevanta effektmått använts för att utvärdera slutresultaten.

Vi tror att EBM har blivit ett så etablerat begrepp bland läsarna att det inte längre behöver presenteras under en egen vinjett. Vi kommer givetvis att även i fortsättningen publicera artiklar som berör den viktiga skärningspunkten mellan klinisk vetenskap och klinisk praxis men nu under vinjetten »Systematiska översikter«. Vi välkomnar fler bidrag som kritiskt granskar välgjorda systematiska litteratursammanställningar.

Mats Eliasson

medicinsk redaktör (evidensbaserad medicin)

Josef Milerad

medicinsk chefredaktör