

Evidensbaserad läkemedelsvärdering

Hans Liedholm, Läkemedelskommittén, Blekinge

Inledning

För en läkare är det viktigt att vara uppdaterad om de medicinska framstegen vad avser sjukdomslära, diagnostik och terapi. "Att arbeta enligt evidensbaserad medicin (EBM) är att följa den kunskap som är insamlad, definierad och värderad och funnen acceptabel enligt vissa normer". EBM kallas även för "kunskapsbaserad medicin". EBM betraktas av vissa som ett hot mot de självständiga besluten, men samtidigt som otillräcklig ute "i verkligheten". Det är som om vetenskaplig metodik skulle stå i ett motsatsförhållande till den egna omdömesförmågan. Detta är fel. EBM fritar oss inte på något sätt från ansvaret att fatta självständiga beslut. Den ger oss bara ett bättre beslutsunderlag.

Kunskap och metodik

En översikt av vetenskapliga rapporter i medicinska tidskrifter visar att det behövs två ting för att förstå dess innehåll. Det ena är kunskap om de förhållanden som undersöks (sjukdomen, terapierna etc) och det andra är en kunskap om de vetenskapliga metoder som använts och den därpå följande tolkningen av resultaten – särskilt kunskap om medicinsk statistik.

Svagheter i medicinsk forskning

Vetenskapliga rapporter måste läsas mycket kritiskt, läsaren får inte ta några resultat för givna. Ändamålet med kritisk läs-

ning är att upptäcka om metoderna kan ge användbar information och om slutsatserna som författarna dragit följer av studieresultaten. Rapportering om medicinsk forskning är en svår uppgift. Därför behövs det också bra anvisningar, råd och mycket träning i att läsa och kritiskt bedöma sådan.

Medicinska litteraturen

Uppgifter om ny kunskap i den medicinska världen rapporteras snabbt. Redan för drygt 20 år sedan angavs antalet publicerade biomedicinska artiklar till > 2 miljoner/år, dvs ca 5 500/dag. En annan uppgift anger att > 17 000 böcker publiceras/år. Idag har det blivit allt vanligare att de traditionella tidskrifterna publicerar angelägna artiklar på Internet innan de kommer ut i pappersversion, samtidigt som antalet enbart Internetbaserade tidskrifter ökar. Tidskrifter eller artiklar som har epitetet "open access", dvs läsbara och nedladdningsbara från Internet utan prenumeration ökar också.

Epidemiologiska studier vs RCT

Eftersom vi får så mycket information är det viktigt att vara uppmärksam på vad för slags data som presenteras. Det framgår sällan av rubrikerna. En del handlar om vad någon auktoritet skriver om, t ex platsen för en viss terapi, annat om epidemiologiska studier och annat om väl genomförda, kontrollerade studier (randomised controlled trials, RCT).

Det första tillhör oftast den s k "åsiktslitteraturen", som har lågt värde på EBM-skalan över trovärdighet (se s 1096).

Det andra tillhör undersökningar som har viktiga begränsningar. Många års data från patientkohorter pekade på att postmenopausal behandling med hormonterapi hade goda effekter på kvinnors hälsa, bl a beträffande hjärt-kärlsjukdom. Under loppet av bara några år kunde RCT tvärtom påvisa behandlingens skadliga effekter.

Ytterligare ett exempel förtjänar att omnämnas. De senaste 25 åren har epidemiologiska sammanställningar upprepade gånger stärkt synen på att förhöjda nivåer av homocystein innebär en ökad risk för kardiovaskulär sjukdom samt att det finns ett linjärt samband mellan homocysteinnivå och risk. För att undersöka effekterna av terapi som sänker homocystein (folsyra, vitamin B₁₂ och B₆) på kardiovaskulära händelser har successivt flera placebokontrollerade studier nu visat att sjukdomsförloppen inte alls påverkas, trots att homocystein sänkts.

Detta har tex visats i en sekundärpreventiv 2-årsstudie av patienter med TIA och stroke (1) och en 3-årsstudie av patienter med färsk hjärtinfarkt (2) samt en 5-årsstudie av patienter med hjärt-kärlsjukdom eller diabetes (3). I den första studien utgjordes de undersökta patienterna bara av sådana som från start hade homocysteinvärden i den högsta kvartilen. Ingen vitaminbehandling var bättre än placebo, i en av studierna fanns också tolkningen att den mest intensiva behandlingen to m kunde öka den skadliga effekten man var ute efter att förebygga.

Det förefaller som om man hela tiden bara påverkat en surrogatvariabel. Epidemiologiska data stämde inte med den kliniska verkligheten!

Låt oss begränsa oss till att säga att den epidemiologiska studiens beräkningar bygger på datainsamling från grupper av patienter, där olika registrerade egenskaper ställs mot varandra med diverse justeringsmetoder för basala olikheter. Här finns ett inbyggt metod- och trovärdighetsproblem. Det man registrerat är det som kan användas och man kan aldrig justera för det man inte registrerat. Epidemiologiska data kan antyda, men aldrig klargöra orsakssamband.

Det tredje, RCT, utgörs av experimentella data som kan göra orsakssamband bevisba-

ra. Det är därför RCT anses ligga högst på skalan av bevisvärde. Exempel på inskränkningar med RCT är brist på generaliserbarhet, resultatet gäller bara de undersökta, den population som undersöktes (bakgrund, sjukdomsgrad, kön, ålder) och den tidsrymd studien varade. Detta betyder samtidigt att all kunskapsinsamling inte kan bestå av RCT, både denna metod och den epidemiologiska har sina fördelar och sina begränsningar. Det viktigaste för läsaren är att förstå vad dessa består av.

Publikationsformen

Tidskrifter som använder sig av oberoende referenter för granskning av insända manus (referee-system) har störst vetenskaplig trovärdighet, liksom de artiklar som publiceras däri. De flesta större tidskrifter som ges ut av internationella förlag följer en likartad arbetsprocess och presentationsform och har skapat gemensamma rekommendationer för krav på artikelförfattarna och manuskripten (se CONSORT, s 1089). En artikel publicerad i en referentbedömd tidskrift är dock inte någon absolut garanti för kvaliteten på det vetenskapliga innehållet (se nedan om fusk)!

Vad kan man lite på?

Det mesta som presenteras i icke referentbedömda artiklar ska betraktas med skepsis. Man bör betrakta de uppgifter som förekommer i "data on file", "abstract" eller i producenternas inköpta supplement till någon känd tidskrift som obevisade eller i varje fall otillräckliga.

Det vilar på författarna (och i viss mån på tidskriftsredaktörerna) att visa upp vad som verkligen är gjort i en studie. Bevisbördan ligger alltså hos författarna. En känd engelsk statistiker ger i en bok ett tänkvärd råd: "Har de inte skrivit det – utgå då från att de inte har gjort det heller!"

Ett annat och förhoppningsvis sällsynt moment som verkligen är svårt att genomskåda är forskningsfusk. År 2005 blev en korean respektive en norrman mycket uppmärksammade för sina forskningsresultat de fått publicerade i några kända tidskrifter, men senare avslöjades att de hade fuskat – helt enkelt hittat på data. Det är i den moderna, snabba kunskapsutvecklingens tide-

varv lätt att glömma bort att en av vetenskapens grundvalar är att resultat ska vara reproducerbara.

Bedömning

En enkel metod för att bedöma en klinisk läkemedelsstudie består av tre punkter. Dessa ska löpa som en röd tråd genom artikeln

- Mål, primär effektvariabel
- Statistik
- Resultat

Viktiga hållpunkter i en artikel

Mål, primär effektvariabel

Vilket var det primära målet med prövningen? Målet ska vara tydligt och väl preciserat. Det anges nästan alltid i slutet av inledningen av artikeln eller i avsnittet där de statistiska metoderna beskrivs. Effektvariabeln bör helst vara enkel och lättförståelig – och av kliniskt intresse – för att räknas. Fortfarande publiceras många artiklar där man inte hittar någon preciserad primär frågeställning.

Syfte

Syftet med en klinisk läkemedelsprövning är, att med utgångspunkt från ett begränsat antal väldefinierade patienter försöka påvisa en skillnad eller (allt oftare) att det nya medlet eller metoden inte är sämre än det man jämför med (s k non-inferiority study).

Avsikten ska givetvis alltid definieras före studiestart. Patientmaterialet måste av etiska och praktiska skäl vara begränsat. Metoderna i den kliniska läkemedelsprövningen ska så mycket som möjligt försöka kontrollera slumpen och preferenser för vissa slutsatser (bias).

Studietyper

Bestäm om studien är en "handhavandestudie" (management study) eller en "förklarande studie" (explanatory study). I det första fallet – vanligast för klinisk läkemedelsprövning – undersöker man vilken medicinsk ändamålsenlighet behandlingen har. I en studie som försöker visa att en behandling är bättre än en annan – en överlägsenhetsstudie – ska alla patienter räknas med vid

"Ny" sammansatt effektvariabel!

I en studie av pioglitazon (Actos) för att eventuellt minska hjärt-kärlsjuklighet och död hos typ 2-diabetiker med hög risk, var huvudresultatet – enligt den primära effektvariabeln – att tillägg av pioglitazon till den antidiabetiska terapin inte var bättre än placebo, naturligtvis en missräkning. Den primära effektvariabeln var en sammansatt effektvariabel som bestod av sju olika kliniska händelser som patienterna kunde drabbas av.

Författarna skapade emellertid en ny sammansatt effektvariabel genom att ta bort fyra av de ingående händelserna. De fann då att pioglitazon var bättre än placebo för de resterande tre. Denna effektvariabel fanns inte med i ursprungsprotokollet till studien. Studien, som publicerades i Lancet, och innehöll den nya tolkningen, var ett klart våld mot CONSORT-principen (se nedan) (4).

bedömning av utfallet. Notera således att en analys enligt "intention to treat" (ITT) räknar in alla i ursprungsgrupperna, oavsett hur de lyckats genomföra studien – "så här skulle det se ut om denna behandling används på vilka patienter som helst" (dvs sådana som kunnat inkluderas i studien). Beräkning och bedömning sker då enligt ITT. Denna analys ska alltid presenteras.

I det andra fallet (förklarande studie) undersöker man hur en behandling egentligen fungerar. Här ska bara resultat räknas från dem som verkligen genomfört studien på ett korrekt sätt. Ett typexempel är en laboratorieundersökning av ett läkemedels kinetik. I ett sådant fall blir den individ som inte kan genomföra studien ersatt av en annan. Analysen sker då enligt principen "per protocol" (PP).

Det är lätt att missa skillnaderna mellan de två studietyperna och tro att resultaten av en vanlig prövning tillhör den andra typen i stället för den första. Om emellertid båda analyserna presenteras kan läsarna själva värdera resultaten.

Statistik

Här ska skälen till dimensionering av studien (antal patienter/försökspersoner) anges och därmed de statistiska förutsättningar-

na för att kunna påvisa skillnader/likheter i den primära effektvariabeln. Hur stor är den skillnad man vill försöka visa mellan t ex två behandlingar? Läsaren ska genast fråga sig om den skillnad man påvisat är kliniskt intressant.

Med vilken precision har man kunnat visa skillnaden, eller med andra ord hur brett är konfidensintervallet för skillnaden i relation till skillnadens storlek? Fortfarande publiceras många artiklar som inte redovisar preciserade statistiska krav på studiens utformning. Under rubriken hör även de metoder som använts för statistiska beräkningar av materialet – metodval om data inte är normalfördelade, om justering gjorts för multipla jämförelser, vilka typer av bortfallsanalyser som gjorts etc.

Resultat

Huvudresultat

Här binds resultatet ihop med den primära frågeställningen och dess statistiska förutsättningar och data redovisas.

Ingen skillnad – inte detsamma som likhet

Efter ett resultat, där den skillnad man hoppats kunna påvisa inte inträffade, försöker ibland författarna (och sponsorerna) hävda att behandlingarna är lika bra. Statistiskt går det dock aldrig att visa att två terapier är lika bra. Däremot kan man visa att skillnaden inte är större än vad som definieras som kliniskt relevant.

Ett exempel på en sådan misstolkning var beskrivningen av den s k NORDIL-studien (5). I denna gällde förutsättningarna att försöka se att den diltiazembaserade terapin gav en 20-procentig skillnad (läs förbättring) i utfallet för den primära, sammansatta effektvariabeln (fatal och icke-fatal stroke, fatal och icke-fatal infarkt samt kardiovaskulär död av andra orsaker). När detta resultat inte uppnåddes hävdade man i stället att "diltiazembaserad" terapi av högt blodtryck var lika bra som behandling baserad på "traditionell" terapi. Den korrekta beskrivningen är att man inte lyckats visa någon skillnad.

"En bild säger mer än tusen ord"

Grafiska presentationer av en studie kan vara allt från upplysande till direkt bedräg-

liga. Huvudprincipen är att data i bild kompletterar tabeller och text. Inte sällan får grafer representera hela studieresultatet. Det är ofta en förenkling – kliniska läkemedelsprövningar är mer komplicerade än så. Se upp med vanligt förekommande suggestiva presentationer med "manipulerad" skalstorlek, avskurna axlar i koordinatsystem, data som inte är normalfördelade, avsaknad av spridningsmått, etc.

Sekundära resultat

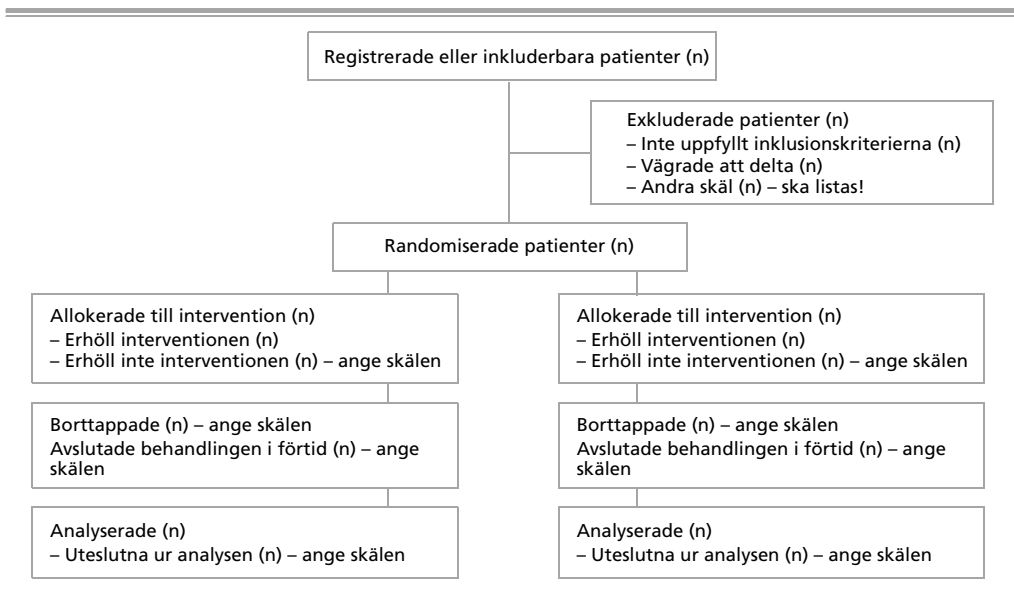
Vid sidan av den primära frågeställningen presenteras ofta resultaten av olika sekundära frågeställningar och resultat från olika undergrupper. Vilka utförda beräkningar är redovisade och vilka saknas? En frågeställning av betydelse kan lätt skapas av en relation som egentligen bara är ett slumpfynd. Sekundära resultat är nästan bara hypotesgenererande och måste undersökas i nya studier med en ny primär frågeställning.

Exempel: När den första jämförande studien av en angiotensin II-antagonist (losartan) och en ACE-hämmare (kaptopril) gjordes på patienter med hjärtsvikt fann man en lägre mortalitet hos dem som behandlats med losartan. Resultatet byggde på en sekundär frågeställning, varför studien gjordes om med ett större patientmaterial och med mortalitet som primär effektvariabel. Man fann då ingen skillnad mellan behandlingarna (6).

Det måste starkt betonas att slumpens inflytande är stort. Man bör betrakta en studie som det den verkligen är – ett stickprov!

CONSORT – att redovisa en klinisk prövning

CONSORT (Consolidated Standards of Reporting Trials) innehåller standardiserade metoder med en checklista och ett flödeschema, vars avsikt är att förbättra rapportkvaliteten hos en RCT. CONSORT är en överenskommelse som har undertecknats av många medicinska tidskrifter. Flödesschemat ger också läsarna en tydlig bild över alla studiedeltagarna i rapporten – från tiden före de blev randomiserade tills de avslutat studien. Den som vill undersöka



Figur 1. Ett typiskt CONSORT-diagram

data i studien kan med flödesschemats hjälp bättre värdera och förstå studiens validitet. I Figur 1 finns ett typiskt CONSORT-diagram.

Vilka var egentligen med i studien?

Tyvärr saknar många publikationer fortfarande detta flödesschema. En annan, inte ovanlig brist, är att det första steget på redovisningen saknas. Många gånger beskriver artikeln bara att man hade ett visst antal patienter med den och den diagnosen och så randomiserades de till den ena eller den andra terapin. Var kom patienterna ifrån – hur många var de från början?

Några exempel: I två svenska studier över behandling av hypertoni hos äldre (STOP Hypertension) angavs bara att det var ett visst antal patienter från studiens start – inte hur de hade tagits ur den svenska primärvårdens patientflöde av äldre hypertoniker (7,8). Var de representativa? Vilka var inte med?

I en välkänd studie över en ACE-hämmare (ramipril) effekter på patienter med kardiovaskulära riskfaktorer (HOPE-studien) fanns inte heller motsvarande bakgrundsuppgifter (9).

I en omtalad jämförande studie av klopidogrel och ASA till patienter med kardiovaskulära sjukdomar (CAPRIE) medgav författarna att de inte visste hur stor ursprungspopulationen var, men väl hur många som randomiserats (10). Även när det finns uppgifter om den screenade populationen så ska man betänka att även denna är selekterad via olika mekanismer.

Information om rekryteringen är således viktig. I en stor studie över effekt av alendronsyra mot osteoporos (se s 1092) framgick det tydligt av publikationen att man för att rekrytera äldre kvinnor med möjlig osteoporos tvingades annonsera efter dem – i det här fallet skickade man ut ca 1 miljon brev för att hitta intresserade kvinnor i de lämpliga åldrarna.

Viktig är också bedömningen av validiteten i studier som medfört stora avbrottsfrekvenser. Exempel på sådana studier är undersökningar av läkemedel mot schizofreni, demens, KOL, alkoholism och niktinberoende.

Validitet

Tillförlitligheten – validiteten – av en undersökning bör bedömas då studieresultat och dess överförbarhet undersöks. Den interna

validiteten är tillförlitligheten i själva undersökningen, t ex noggrannheten i metoderna och studiens genomförande. Den externa validiteten gäller t ex hur resultaten kan överföras till andra grupper med samma sjukdom, dvs hur generaliserbara de är. Hög intern validitet innebär inte automatiskt hög extern validitet, t ex då undersökningen är noggrann men gäller en starkt selekterad patientgrupp. Eftersom en studie med otillfredsställande intern validitet aldrig kan ha en hög extern validitet är det viktigt att särskilt granska metodologin i studien.

Gradering av evidens

Undersökningsresultat (t ex från läkemedelsstudier) har olika tyngd och evidens kan därför behöva grupperas efter sin styrka i någon form av grupper (grader). Ett sätt – dock inte okontroversiellt – utgår från en gradering uppifrån ("bäst") av systematiska översikter av flera väl designade RCT ner till ("sämst") uppfattningar av s k "respekterade auktoriteter". Se mer i texten i föregående utgåva av Läkemedelsboken, LB 2007/2008 (finns även på nätet).

Evidens tolkas olika

Medicinska statistiker har upprepade gånger påpekat att medicinare ofta använder statistik på ett felaktigt sätt, att de ofta inte förstår vilka statistiska metoder som bör användas och att de saknar kunskap om metoder som tillämpas i undersökningar. Fakta som presenteras på olika sätt uppfattas också olika.

Läkare uppfattar överlag resultat som beskrivs i termer av "relativ riskreduktion" mycket positivt. I en undersökning svarade således 49% av 235 läkare att de var mer benägna att skriva ut lipidsänkare om de fick veta att behandlingen ledde till 24% relativ riskreduktion än om resultatet beskrevs som en absolut riskreduktion på 0,4% – trots att beräkningarna gällde samma undersökning.

En likartad gradering gör patienter. I en studie svarade 88% av de tillfrågade patienterna att de var villiga att ta ett läkemedel som gav 34% relativ riskreduktion av hjärt-

attacker mot 31% om de fick veta att 71 patienter måste behandlas i 5 år för att en patient skulle ha nytta av behandlingen.

Hur många måste behandlas?

NNT, ett exempel

Det antal man måste behandla med ett läkemedel för att undvika en enda händelse (t ex stroke) kan uttryckas som NNT – number needed to treat. Detta accepteras som ett meningsfullt uttryck för att beskriva ett läkemedels fördelar mot en jämförelse, t ex placebo. Sammanfattningsvis är $NNT = \text{det inverterade värdet av den absoluta riskminskningen, på engelska } ARR (\text{absolute risk reduction}) = 1/ARR.$

Beräkning av NNT – ett exempel

I en randomiserad (fiktiv) studie med 100 patienter fick 50 det gamla preparatet Obsolet Retard (OR) och 50 det nya Aktivén Forte (AF). Ändamålet med studien var att se om AF minskar insjuknandet i infarkt. Studien pågick i 3 år. Resultatet vid studiens slut framgår nedan.

Terapi	Infarkt		Summa patienter	Risk för infarkt
	Ja	Nej		
Obsolet Retard	20	30	50	0,4
Aktivén Forte	5	45	50	0,1

- Om vi behandlar 50 patienter med OR får 20 infarkt. Risk för infarkt med OR: $20/50=0,4$ (=40%)
- Om vi behandlar 50 patienter med AF får 5 infarkt. Risk för infarkt med AF: $5/50=0,1$ (=10%)
- Relativ risk för infarkt med AF: $0,1/0,4=0,25$ (=25%)
- Relativ minskning av risken med AF: $100-25 = 75\%$
- Absolut riskminskning (ARR): $0,4-0,1=0,3$ (=30%)

$$NNT = 1/ARR = 1/(0,4-0,1) = 1/0,3 = 3,33 \approx 4$$

(hela tal anges). Vi kan förhindra att ytterligare 1 patient får en infarkt under 3 år om vi behandlar 4 patienter med AF i stället för OR.

Vilken är evidensen? – Exempel 1

I en studie, som publicerades 1996, önskade man påvisa en minskning av risken för frakturer med ett osteoporosmedel av bisfosfonattyp (alendronsyra; Fosamax) i jämförelse med placebo (11). Studien gjordes på postmenopausala kvinnor som hade nedsatt bentäthet och som haft 1 kotfraktur. Den planerade behandlingstiden var 3 år. Vid studiens start randomiserades 1 022 kvinnor till aktiv behandling med bisfosfonat och 1 005 till placebo.

Huvudresultatet i studien var att incidensen av kliniska frakturer på alla förutbestämda lokaliseringar minskade signifikant (primär effektvariabel). Multipla kotfrakturer minskade också signifikant. Höftfrakturer – den mest kostsamma frakturtypen för sjukvården och samhället i övrigt – uppstod hos 11 patienter (1,076%) som fått aktiv behandling och hos 22 patienter (2,189%) som fått placebo. Skillnaden som uppmättes var statistiskt signifikant, p-värdet angavs till 0,047.

Hur bra är behandlingen med bisfosfonat enligt denna studie?

P-värdet

Resultatbudskapen i medicinska artiklar cirklar mycket kring p-värden (p står för sannolikhet – probability). Vid statistiska jämförelser är alltid grundhypotesen att det inte finns någon skillnad mellan behandlingsgrupperna, den s k noll-hypotesen. Det uppnådda resultatet är ett av flera man skulle kunna få om studien gjordes om. Med likadana patientgrupper skulle en del visa en större skillnad, andra en mindre skillnad. Man skulle få en viss spridning av skillnaderna beroende på slumpen.

I alendronsyrastudien blev p-värdet 0,047. Sannolikheten att få vad man har fått eller att få något ännu extremare om noll-hypotesen är sann är 0,047. Vilken slutsats man ska dra av det uppnådda resultatet beror egentligen på betraktaren, som själv ska avgöra frågan. Antingen beror skillnaden på slumpen (men är ovanlig) eller så är resultatet ett tecken på att det finns en faktisk effekt av alendronsyra.

Sannolikhetsgränsen sätts oftast vid 5%, vilket bara är en konvention. Men om man har bestämt sig för 5% som gräns är man ock-

så överens om att ett lägre p-värde ($p < 0,05$) betyder att man ska förkasta noll-hypotesen. I ett sådant fall tror vi att grupperna är olika. I var 20:e studie med ett sådant utfall finns dock ingen verklig skillnad (falskt positiv) utan den betingas av slumpen.

P-värdet är sannolikheten att få ett visst utfall och säger heller inget om skillnaden i absoluta tal. Samma numeriska p-värde kan erhållas om man observerat en liten skillnad mellan mycket stora grupper (statistisk signifikans utan klinisk signifikans) eller en stor skillnad mellan mindre grupper.

Relativ risk (RR)

I exemplet blev utfallet av bisfosfonatbehandlingen (noga räknat) 1,076% höftfrakturer och av placebobehandlingen 2,189% höftfrakturer och förhållandet dem emellan $1,076/2,189 = 0,491$. Den relativa risken för höftfraktur vid bisfosfonatbehandling kan sägas vara 49%.

Relativ riskreduktion (RRR)

En 49-procentig relativ risk för höftfraktur betyder att den relativa risken (RR) för fraktur reducerats med 51%. Man kan nu hävda att behandlingen reducerat antalet höftfrakturer med 51%. Siffran ser imponerande ut, men beror på relationen mellan grupperna. Procentangivelsen lurar oss lätt.

Absolut riskreduktion (ARR)

Den absoluta riskreduktionen är förstås viktigare. I exemplet är den 2,189% minus 1,076%, vilket blir $1,113\% = 0,01113$ (med ett konfidensintervall från 0,00011 till 0,02327). Talet förefaller inte så imponerande nu.

NNT – Number needed to treat

Vi behöver emellertid översätta resultatet till mer lätthanterliga siffror. I vårt exempel är NNT $1/0,01113$, dvs 90 (vid exakt beräkning) och detta tal imponerar ännu mindre. För att förhindra ett fall av höftfraktur med denna bisfosfonat måste vi således behandla 90 patienter under i genomsnitt 3 år med läkemedlet (tillsammans med kalcium).

Grafsk beräkning av NNT

En snabb uppfattning av NNT-värdet kan man få genom att använda ett nomogram

sedan man har erhållit värden för absolut risk och relativ riskminskning, se Läkemedelsboken 2007/2008 (finns även på nätet).

Konfidensintervall

Vi kan inte bara nöja oss med resultatet från den genomförda undersökningen. Vi behöver veta hur pass osäkert det angivna NNT-värdet är. Precis som för p-värdet så kommer vi att få andra NNT-värden om vi genomför fler undersökningar med samma förutsättningar. Genom att redovisa ett konfidensintervall kring NNT-värdet kan vi förmedla den bilden. Men den nedre och övre gränsen för konfidensintervallet placerar sig inte symmetriskt runt NNT-värdet och överraskar nog.

Av hävd anges oftast det 95-procentiga konfidensintervallet, vilket betyder att det har 95% chans att falla så att det täcker det sanna värdet för NNT. Man har alltså 95% konfidens (förtroende) för intervallet.

I vårt exempel kan vi med hjälp av ett behändigt dataprogram räkna ut att konfidensintervallet ligger mellan 43 och 8 949 patienter.

Vilken beskrivning tilltalar mest?

Uppgifterna om evidensen i alendronsyra-studien kan nu jämföras. Vilken ska vi välja för att bäst beskriva studieresultatet?

- Höftfrakturer minskas signifikant av alendronsyra, $p = 0,047$.
- Alendronsyra minskar höftfrakturincidensen med 51%.
- Den absoluta riskreduktionen för höftfraktur är 1,1%.
- För att förhindra en höftfraktur måste 90 patienter behandlas med alendronsyra i 3 år.

Siffran 90 är lite osäker – antalet ligger någonstans mellan 43 och 8 949 patienter (95% konfidensintervall).

Underlag för kostnadsberäkning

Nu blir det lätt att beräkna kostnaden för läkemedelsinsatsen. Om vi måste behandla 90 patienter i 3 år för att förhindra en höftfraktur uppgår kostnaden till 270 x årskostnaden för en behandling, dvs 270 x 4 000

Effektbeskrivningar

- Statistisk signifikans
- RRR, relativ riskreduktion
- ARR, absolut riskreduktion
- NNT med konfidensintervall

kronor (dåtida prissättning, avrundade tal) = 1,08 miljoner. Detta värde är lite osäkert – den sanna läkemedelskostnaden ligger med 95% sannolikhet mellan 0,5 och 107 miljoner kronor.

Värdet av insatsen kan bara bedömas efter en omfattande och ingående analys där man tagit hänsyn till alla direkta och indirekta kostnader för frakturer, hur läkemedelsbehandlingen tolereras i en mer allmän patientgrupp än i studiepopulationen, hur mycket biverkningarna kostar i lidande och pengar, vad långtidseffekterna egentligen består av etc. Se vidare kapitlet Fallgropar vid hälsoekonomisk läkemedelsvärdering LB 2007/2008 (finns även på nätet).

Vilken är evidensen? – Exempel 2

Följande data är hämtade från ett abstrakt, men kan ändå tjäna som diskussionsunderlag för olika sätt att uttrycka storleken av ett utfall.

I en rapport som bedömt frekvensen allvarliga hypoglykemiska händelser i en sammanslagning av fyra 1-årsstudier jämfördes ett nytt antidiabetesmedel (repaglinid) med tre andra diabetesmedel av SU-typ (glibenklamid, glipizid och gliklazid (inte godkänt i Sverige)).

Man fann att frekvensen (händelser/år) av hypoglykemi hos repaglinidbehandlade (sammanlagt antal patienter = 761) var 1,31%, medan den hos de övriga (sammanslaget antal patienter oaktat att det var tre olika jämförelsepreparat = 367) var 3,27%. Studiernas uppläggning angavs ha varit likartad och alla patienter hade uppmanats att inta regelbundna måltider.

Texten i den aktuella informationen lyder ungefär så här: ”I jämförelse med repaglinid ger SU-preparat som givits till typ 2-diabetiker en 2,5-faldig ökning av symtomatisk hypoglykemi.”

Flera kommentarer

Man har jämfört ett medel med tre andra. Genom att "klumpa ihop" resultat från 3 preparat kan man ha misskrediterat ett eller två av dessa med ett sämre läkemedel.

Oavsett lämpligheten i den bakomliggande databearbetningen i sammanställningen, kan vi beskriva data på ett annat vis. Om vi får förutsätta att patienter som råkat ut för hypoglykemi bara räknas en gång i studien, består antalet patienter som fått hypoglykemi i repaglinidgruppen av 10 patienter, hos de övriga terapierna av 12 patienter. Räknar vi ut NNT för att förhindra en hypoglykemi under 1 år är NNT = 51.

Man behöver således behandla 51 patienter med repaglinid i 1 år för att undvika en hypoglykemi jämfört med de övriga SU-preparaten. Det 95-procentiga konfidensintervallet för NNT ligger mellan 23 och 460 individer.

Underlag för kostnadsberäkning

Om vi antar att dosen för repaglinid i studien var 1 mg 3 gånger/dag med en (dåtida) årskostnad på 2 838 kronor och dosen för SU-preparatet glibenklamid var 3,5 mg 2 gånger/dag med en årskostnad på 675 kronor blir merkostnaden för repaglinid 2 163 kronor/patient. Att förhindra en hypoglykemi kostar för NNT = 51 patienter drygt 110 000 kronor. Då man tar hänsyn till konfidensintervallet är merkostnaden mellan knappt 50 000 och knappt 1 miljon kronor/år för varje fall av hypoglykemi som undvikits. Om det inte finns någon egentlig effektskillnad mellan preparaten i övrigt kan man fundera på den ökade marginalkostnaden.

I ovanstående liksom i föregående exempel bygger uppgifterna bara på direkta läkemedelskostnader, endast en av komponenterna som TLV (Tandvårds- och läkemedelsförmånsverket, tidigare LFN, Läkemedelsförmånsnämnden) baserar sina beslut på.

Behandling utan evidens på effektivitet

NNT-diskussionen ovan visar att vi ofta har en mycket optimistisk syn på olika terapiers effektivitet. Även om vi behandlar på allra bästa sätt med ett läkemedel, som t ex ska

fungera förebyggande (t ex ASA), kommer många patienter ändå att drabbas av det vi försöker förebygga! Ett behandlingsmisslyckande (t ex stroke trots ASA-behandling) leder ofta till slutsatsen att behandlingen nu behöver bytas mot ett "modernare" medel (t ex klopidogrel, Plavix). Patienten ser sig sannolikt bättre omhändertagen av åtgärden, men vetenskapligt stöd för behandlingsändringen saknas. Det är ändå troligt att den behandlande läkaren uppfattar bytet som effektivt då det kanske dröjer länge innan en ny stroke inträffar. Bytet till den nya behandlingen uppfattas ske enligt "beprövad erfarenhet" men saknar evidens. Se vidare i kapitlet Cerebrovaskulära sjukdomar, s. 317.

NNT-exempel

I Tabell 1 finns en lista över några NNT-värden med 95% konfidensintervall. Data är hämtade från några beskrivningar av Bandolier i Oxford (Evidence-based health care) och Mount Sinai Hospital, Centre for Evidence-Based Medicine.

Tabellen avser inte att ge en sammanfattande bild av genomsnittseffekten av de olika behandlingarna. Därför kan de uppfattas vara i konflikt med andra uppgifter i LB och annorstädes. Resultaten gäller behandlingar där data kommer från enstaka, publicerade studier.

Resultaten kan naturligtvis inte jämföras rakt över eftersom studierna behandlar helt olika patientgrupper. För att rätt kunna tolka resultaten måste man ha tillgång till många andra uppgifter i studierna, t ex kriterierna för de patienter som deltagit i undersökningen. På så vis kan man hävda att "varje studie bär sina egna NNT-värden".

Det är också viktigt att beakta skillnader i samtidig medicinsk behandling som patienterna har. Genom en övergång till andra terapitraditioner kommer patientpopulationerna inte att vara jämförbara med varandra – inom landet eller mellan länder – om tillräcklig tid förlöper mellan de olika studierna. NNT-värden är då ännu svårare att ta fram och jämföra!

Ett problem är förstås att uppgifter av det här slaget inte finns till hands då man behöver dem. En möjlighet är att ta in NNT-data

Tabell 1. Exempel på NNT-värden och deras 95-procentiga konfidensintervall

Åkomma	Behandling	Jämförelse-behandling	Behandlingstid	Utfall	NNT	95% konfidensintervall
Alzheimers sjukdom	Ginkgo biloba	Placebo	1 år	4 poängs förbättring i ADAS-Cog	8	4–67
Impotens (erektil dysfunktion)	Alprostadil (givet i uretra)	Placebo	3 månader	Grad av erektion som kunnat leda till genomförande av samlag	2	2–3
Hundbett	Antibiotika	Placebo	Korttids-behandling	Hindra sårinfektion	16	9–92
Huvudlöss	Insekticida medel	Placebo	1 dos	Frånvaro av viabla löss eller lusägg efter 14 dagar	2	1–2
Herpes zoster	Aciklovir	Placebo	5–10 dagar	Smärta efter 1/2 år	–	Inte signifikant
Osteoporos	Kalcium och D-vitamin	Placebo	3 år	Förebygga fraktur	20	13–57
Lätt eller stark smärta efter operation	Paracetamol 1 g i enkeldos	Placebo	1 dos	Minst 50% smärtlindring	4	3–5
Smärtande, stel axel	Lokal steroidinjektion	Fysikalisk terapi	6 veckor	Behandlingsframgång efter 1 år	3	2–8
Akut mediaotit	Xylitol-tuggummi	Vanligt tuggummi	3 månader	Minskad incidens av akut otit	8	4–53
Kataraktoperation av det andra ögat	Operation	Kontroll	6 månader	Minskade svårigheter att läsa normal text	4	3–6
Peptiskt sår	Trippelterapi	H ₂ -blockerare	6–10 veckor	Ulcusläkning efter 6–10 veckor	5	4–7

i rekommendationslistor från läkemedelskommittéer. Vid diskussioner där en behandling anges vara bättre än en annan bör man samtidigt tvingas att svara på hur stort NNT-värdet är. Ännu mer värdefullt är att ange ett konfidensintervall så att man förstår osäkerheten i uppgifterna!

Risken för din patient?

Hur kan man föra över ett resultat från en studie till en enskild patient du har på mottagningen med samma sjukdom? För ett rimligt svar på den frågan måste du fråga dig följande:

- Skiljer sig din patient från dem som ingick i studien på något avgörande sätt? Ålder, kön, annan samtidig sjukdom, andra simultana läkemedel?

- Kan du bedöma om din patient har samma risk för sjukdom eller komplikation som i den behandlade gruppen?
- Har din patient samma risk för biverkningar som den behandlade gruppen?

Antalet patienter som måste behandlas för att en patient ska ha nytta av behandlingen – NNT – är ett slags genomsnitt för effekt. Patienter med stor risk har en bättre möjlighet att dra nytta av behandlingen; patienter med låg risk har mindre nytta av behandlingen. Har din patient en högre risk ökar möjligheten att din patient kommer att vara ”gynnad” – tvärtom ifall risken är lägre. Kan du bedöma frågorna kan du nu bättre avgöra om du ska acceptera behandlingen!

NNH – Number needed to harm

För att beskriva hur biverkningarna skiljer sig mellan två terapier kan man på motsvarande sätt som för NNT beskriva hur många man måste behandla tills en patient drabbas av en viss negativ effekt, NNH.

NNH – exempel 1

Effekten av klopidogrel (Plavix) som tillägg till ASA vid instabil angina pectoris studerades i en randomiserad, placebokontrollerad studie (CURE). En sammansatt primär effektvariabel var död av kardiovaskulära orsaker samt icke dödlig hjärtinfarkt eller stroke. Behandling gavs i medeltal i 9 månader. I den grupp som fick kombinationsbehandling (n = 6 259) uppnådde 582 patienter (avrundat 9,3%) denna händelse mot 719 patienter (avrundat 11,4%) av dem som enbart fick ASA (n = 6 303). Absolut riskreduktion var 2,1%, dvs NNT blir 48 individer (32–96 individer enligt 95% konfidensintervall).

Behandlingen ledde dock till en ökad frekvens blödningar. Antalet allvarliga blödningar inträffade hos 231 individer med kombinationsbehandlingen (avrundat 3,7%), mot 169 individer som enbart fått ASA (avrundat 2,7%). NNH blir 99 individer (61–251 individer enligt 95% konfidensintervall).

Balansen mellan effekt och biverkningar av behandlingen kan beskrivas på följande sätt. Behandlar man 96 patienter (2 x 48) med klopidogrel i kombination med ASA i medeltal under 9 månader förhindras 2 patienter, utöver vad enbart ASA åstadkommer, att dö av en kardiovaskulär orsak, få en icke dödlig hjärtinfarkt eller stroke. Samtidigt drabbas en patient av en allvarlig blödningsskomplikation.

NNH – exempel 2

Behandling med erythropoietin kan förbättra hemoglobinvärden hos olika patienter med anemi. I en systematisk översikt och metaanalys av effekter av erythropoietinstimulerande läkemedel till patienter med cancerrelaterad anemi som jämfördes med placebo eller annan kontrollbehandling, förbättrades hemoglobinvärdena medan mortalitet och venös tromboembolism ökade efter en behandlingstid på 6 till 52 veckor (12). För tromboemboliska händelser uppgick NNH

- Patienter som deltar i kliniska prövningar skiljer sig ofta på betydande punkter från de som sedan ska få aktuellt läkemedel.
- I "verkligheten" blir effekten av behandlingen ofta mycket sämre.
- Genom balansen av beräkningsresultat av NNT och NNH kan man på ett tidigt stadium få en första, nyanserad bild av hur en terapi kan fungera.

till 37 (24–67 individer enligt 95% konfidensintervall). Denna med flera andra studier har lett till revision av bedömning och användning av erythropoietinläkemedel.

Brist på evidens

På många områden finns det inte tillräckligt med studier för att dra några praktiska slutsatser om vad som är bra eller dåligt. "Det finns ingen evidens." Ibland misstolkas detta som att de metoder som används är felaktiga eller värdelösa. Så är inte fallet – brist på evidens är inte liktydigt med evidens på bristande effekt – "absence of evidence is not evidence of absence".

Genväg till kunskaper? Systematiska översikter och metaanalyser

För att få reda på aktuell, sammanlagd kunskap om en viss terapi vill man gärna läsa någon form av översikt. En sådan översiktsartikel kan i engelskspråkig litteratur vara det som kallas för "review" eller en "systematic review".

Systematisk översikt

En systematisk översikt är en artikel som efter en viss mall refererar och kommenterar alla tillgängliga resultat som belyser ett visst medicinskt problem. Till skillnad från en metaanalys görs inte någon statistisk sammanslagning av de refererade resultaten. Metoden försöker minska sannolikheten för systematiska fel (bias). Eftersom bedömningar görs i flera olika steg i en översikt (ex-

empelvis bestämning av inklusions- och exklusionskriterier, kvalitetsbedömning av ingående studier) är systematiska översikter inte befriade från bias. Ett problem som ofta diskuteras är s k "publication bias", dvs att positiva studier är lättare att publicera än negativa. Upprepade publikationer av samma data och framför allt undanhållande av data från offentliggörande grumlar också bilden och kan medföra att översikter och metaanalyser blir missvisande.

Tyvärr leder ibland systematiska översikter från olika grupper av författare till helt olika resultat, en faktor som påvisats är den bindning författarna har till uppdragsgivaren. Förvirringen hos läsaren blir stor – vem kan man lita på?

Metaanalys

Bättre överblick kan man få om man läser en metaanalys av undersökningar i ämnet. En metaanalys är en metod för att göra en samlad bedömning av ett antal jämförande undersökningar genom att statistiskt sammanföra resultaten. Man gör en omfattande litteratursökning och visar hur och var man sökt. De språkområden och de tidsperioder man undersökt framgår tydligt. Man granskar allt publicerat material och man försöker också få fram opublicerat material.

Man ställer upp kriterier för vilka undersökningar som ska ingå i analysen. Resultatet för varje undersökning redovisas i form av ett jämförande resultatmått (t ex en oddskvot med konfidensintervall) som avslutas med en statistisk sammanslagning av resultaten till ett samlat resultatmått (med konfidensintervall).

Framställningen ger en överblick över tillgängliga resultat och deras samstämmighet (homogenitet). Resultatet ger en uppfattning om huruvida undersökningarna visat att en metod är bättre än en annan.

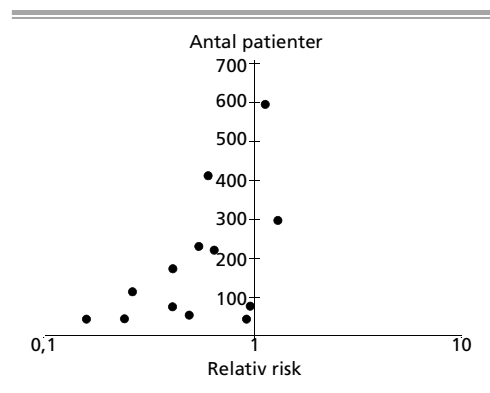
Inte heller metaanalyser är "ofelbara". Det metodologiska arbetet är viktigt och de studier som tas med i analysen avgör utfallet. Baseras analysen på mindre studier kan slutsatsen i metaanalysen vara missvisande. En grafisk metod att belysa betydelsen av de ingående studiernas storlek är att redovisa dem i en s k funnel plot. I en sådan ser man storleken av varje studie (antal patienter i studien) avsatt mot effektstorlek-

en (visas t ex i form av relativ risk). Om det inte finns en "publication bias" förväntar man sig att resultat från mindre studier kommer att fördela sig normalfördelat omkring medelvärdet. Ett konstruerat exempel på en funnel plot visas i Figur 2.

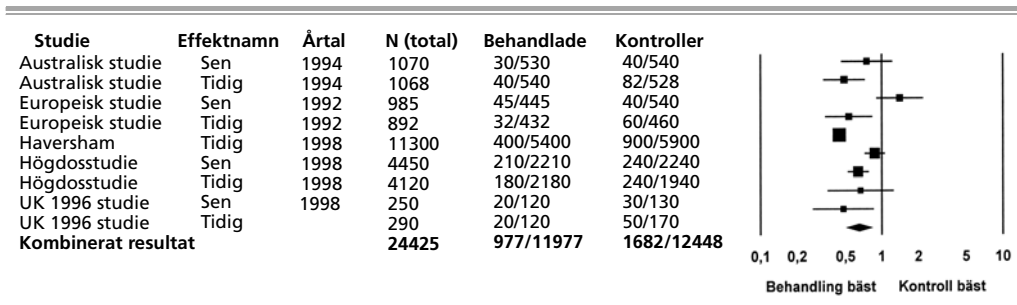
Ett intressant exempel med riskerna av att ha ett selekterat urval av studier att tillgå för en bedömning framgår av en svensk undersökning. I den jämfördes 42 placebokontrollerade studier av 5 antidepressiva medel (SSRI) som skickats in till Läkemedelsverket med den faktiska publikationen av samma studier (13).

Av studier som bedömts positiva för de undersökta substanserna hittades 19 av 21 studier som enstaka publicerade artiklar. För 21 "negativa studier" återfanns endast 6 som enstaka publikationer, resten var inlagda i andra publikationer. Tjugoen studier publicerades minst 2 gånger medan 4 "negativa" studier förblev opublicerade. Flera studier rapporterade dessutom inte data enligt ITT, utan hade en per protocol analys, som därmed gav en mer fördelaktig beskrivning av preparatet.

I och med att alla kliniska prövningar inom EU från och med maj 2004 måste registreras i en databas (EudraCT, kan nås via



Figur 2. Exempel på funnel plot. Varje symbol representerar en studie. Antalet patienter i studierna är avsatta mot den vertikala axeln. Effekten av två behandlingar som jämförts uttrycks som relativ risk – Vid värdet "1" är behandlingsresultaten lika. Bilden belyser problemet med att basera slutsatser på flera mindre studier.



Figur 3. Exempel på en metaanalys. Den understa raden visar den kombinerade effekten från olika studier, t ex dödsfall, ny infarkt eller annan händelse. Den horisontella axeln är logaritmisk. Vid värdet "1", som uttrycker relativ risk eller oddskvot, är resultaten i grupperna lika. Symbolernas storlek är proportionella till studiernas storlek. De horisontella strecken utgör omfånget av ett 95% konfidensintervall för värdet.

<https://eudract.emea.europa.eu/index.html>) kan kontrollen av "publication bias" bli mycket enklare och bättre.

- Vad presentationer av metaanalyser ska innehålla har arbetats fram i en arbetsgrupp och benämns QUOROM (Quality of Reporting of Meta-analyses, jfr CONSORT).
- Motsvarigheten för presentation av observationsstudier benämns MOOSE (Meta-analysis of Observational Studies).
- Motsvarigheten för kvalitativa studier benämns RATS (BMC Qualitative Research Review Guidelines – RATS).
- Det finns även en standard för studier över diagnostiska tester – STARD (Standards to Improve Reporting of Studies of Diagnostic Tests). Enligt en publicerad översikt från 2006 var ett vanligt fel att studier, i syfte att beskriva nya diagnostiska tester, ofta var kraftigt underdimensionerade (låg power).
- Senast i familjen av anvisningar finns nu förslaget STROBE, som avser att strama upp beskrivningen av observationsstudier.

Den internationella intresseorganisation, som bl a är specialiserad på att genomföra metaanalyser, är Cochrane Collaboration (CC), som kan nås på <www.cochrane.org>.

Studierna som presenteras i Cochrane Library är högt rankade och står för kvalitetsvär-

derad kunskap. Den nackdel man kan se med arbetet i CC är den tid det tar att ta fram materialet för sammanställningarna.

Symbolen för CC är den grafiska återgivningen av resultaten i en metaanalys, en s k "forest plot". Ett exempel på en grafisk redovisning av en metaanalys visas i Figur 3.

Andra källor

Ett bra exempel på hur man kan sammanställa evidensbaserad kunskap är BMJ Clinical Evidence, nedan kallad CE (på Internet och som Handbook). Innehållet och antal ämnesområden i CE utvidgas snabbt och belyser vardagsnära medicinska problem. Tillgången är helt kostnadsfri för bl a engelska allmänläkare. Några svenska landsting har prenumererat på servicen.

En hel del publicerade, systematiska översikter kvalitetsbedöms och kan sökas på Internet hos CRD (Centre for Reviews and Dissemination). Hos CRD kan man även hämta in analyser av hälsoekonomiska bedömningar och andra rapporter från många HTA-organisationer (SBU är en av dessa).

Kunskap är en färskvara

Ett skämtsamt citat går ut på att dagens lärare vet att av det de lär ut är bara hälften "sant", problemet är bara att de inte vet vilken halva. Därför bör man lägga in en brasklapp över presenterade resultat, de må vara i form av systematiska översikter eller metaanalyser. Nya data kan lätt ge en ny bild av "sanningen". Därför bör systematiska översikter ersättas av nya. Cochrane Collabora-

tion ser ständigt över sina slutsatser av meta-analyserna och uppdaterar dessa.

Det finns således varningstecken utsatta för hur mycket man kan lita på sammanställningarna. I den medicinska litteraturen finns exempel på hur metaanalyser som kommit till vissa slutsatser antingen har avfärdats eller bekräftats sedan enstaka, mycket stora RCT genomförts. Exempel är intravenös magnesiumterapi (ISIS-4) respektive streptokinasbehandling vid akut hjärtinfarkt (GISSI-1 och ISIS-2) (14–16).

Evidens och rekommendation

I ett avsnitt i föregående utgåva av Läkemedelsboken 2007/2008 (finns även på nätet) redovisades tre tabeller (Tabell 2–4) med olika rekommendationer för läkemedelsterapi – läsaren hänvisas till dessa i den föregående utgåvan. Sedan dess har GRADE alltmer vunnit acceptans.

GRADE

I CE började man 2007 liksom senare flera medicinska organisationer, i Sverige SBU, Socialstyrelsen och Läkemedelsverket, anamma och introducera en metod för att värdera kvaliteten i publicerat forskningsresultat och bedöma styrkan i rekommendationer. Det bygger på ett poängsättningsystem som utvecklats av arbetsgruppen Grading of Recommendations Assessment, Development and Evaluations (GRADE) (17).

Arbetet i CE avser göra begränsningarna i översikterna så att användarna kan se vilka resultat som har betydelse och för vem som berörs. Tillvägagångssättet i GRADE försöker också att peka på svagheter i evidensgraderingen.

Viktigt är värdering av evidens för enstaka studier men även mellan olika studier och då ta hänsyn till

- metodologiska svagheter inom studierna
- samstämmighet (consistency) i resultaten mellan studierna
- hur generaliserbara forskningsresultaten är för en större patientgrupp
- hur bra eller skadliga behandlingarna visat sig vara.

Alla behandlingsjämförelser poängsätts och presenterar kvalitet av evidensen såsom

- hög
- måttlig
- låg eller
- mycket låg kvalitet.

Våren 2008 fanns beskrivningar av mer än hälften av alla systematiska översikter i CE samt tillkommande uppdateringar och nya ämnen.

I CE har länge funnits bedömningar (även i grafisk symbolform) av huruvida en behandling bedöms som sannolikt bra, ineffektiv eller man inte vet eller inte vet tillräckligt för att värdera effekterna. Dessa bedömningar följs av länkade PubMed abstrakt.

GRADE-systemet avses införas i dokument från Socialstyrelsen och SBU. Cochrane Library och flera tidskrifter har också anslutit sig. Vid en genomgång av AHRQ, som jämförde olika system för evidensgradering för ett par år sedan, befanns GRADE-systemet vara det som bäst uppfyllde kraven bland rankade system.

För de som vill ha enkla riktlinjer kan GRADE-bedömningarna uppfattas som lite avancerade, kanske för att de liknar verkliga förhållanden. Rekommendationer, tex från läkemedelskommittéer, har ofta saknat tydliga beskrivningar av evidens och skäl till rekommendationer av terapi. Är det dags för ett paradigmskifte för beskrivning av bakgrundsrekommendationer?

Var kan jag hitta EBM-uppgifter?

Ett känt nätverk av institutioner som lätt hittas på Internet är Cochrane Collaboration (CC). Man kan prenumerera på deras systematiska översikter via CD-skivor (med uppdateringar) för bruk på en egen dator eller direkt via Internet. Via förlaget Wiley & Son når man nu till Cochrane Library.

Många länder har skaffat nationella licenser för direktaccess – i Sverige via SBU (www.sbu.se).

Nedan följer en lista över några Internet-adresser (sites) som kan vara av värde. Flera adresser innehåller länkar till liknande områden.

- SBU – Statens beredning för utvärdering av medicinsk metodik. SBU-Alert via samma länk.
www.sbu.se
- Rapporter från det internationella nätverket (etablerat 1993) för SBU-liknande organisationer, International Network of Agencies for Health Technology Assessment – INAHTA – hittas via följande länk:
www.inahta.org
- Den kanadensiska motsvarigheten till SBU heter Canadian Agency for Drugs and Technologies in Health – publicerar en stor mängd lättöverskådliga EBM-rapporter.
www.cadth.ca
- The Cochrane Library – nås kostnadsfritt via SBU.
www.sbu.se
- NICE – National Institute for Health and Clinical Excellence. Organisation i England som bedömer medicinska behandlingar. Verksamheten är organiserad under NHS (National Health System).
www.nice.org.uk
- The Centre for Evidence Based Medicine; i Oxford under NHS R&D. Flera praktiska länkar.
www.cebm.net/index.asp
- Centre for evidence based medicine Mount Sinai Hospital – University Health Network. Lär ut att öva och praktisera EBM.
www.cebm.utoronto.ca
- Netting the Evidence – A ScHARR Introduction to Evidence Based Practice on the Internet. Utomordentlig lista över EBM-länkar som produceras av Andrew Booth vid School of Health and Related Research (ScHARR), Universitetet i Sheffield, England.
www.shef.ac.uk/scharr/sections/ir/
- Bandolier – site i Oxford med översikter och kritiska bedömningar av alla slags medicinska åtgärder, inklusive behandling med läkemedel. Nytt nummer varje

månad sedan 1994.

<http://www.jr2.ox.ac.uk/Bandolier/>

- Therapeutics Initiative – Evidence based information on drug therapy, University of British Columbia, Vancouver, Canada. Ger ut tidskriften Therapeutics Letter.
<http://tiubcca.nationprotect.net/en/TherapeuticsLetters>
- Länk till Clinical Evidence:
www.clinicalevidence.com/ceweb/
- Länk till CONSORT.
www.consort-statement.org/
- Flera länkar: MOOSE, QUOROM, STARD och STROBE
www.consort-statement.org/?o=1219
- CRD Databases:
www.york.ac.uk/inst/crd/welcome.htm
- Länk till RATS:
www.biomedcentral.com/info/ifora/rats

Referenser

1. Toole JF, Malinow MR, Chambless LE, Spence JD, Pettigrew LC, Howard VJ, et al. Lowering homocysteine in patients with ischemic stroke to prevent recurrent stroke, myocardial infarction, and death: the Vitamin Intervention for Stroke Prevention (VISP) randomized controlled trial. *JAMA*. 2004;291:565–75. [PubMed ID: 14762035]
2. Bønaa KH, Njølstad I, Ueland PM, Schirmer H, Tverdal A, Steigen T, et al; NORVIT Trial Investigators. Homocysteine lowering and cardiovascular events after acute myocardial infarction. *N Engl J Med*. 2006;354:1578–88. [PubMed ID: 16531614]
3. Lonn E, Yusuf S, Arnold MJ, Sheridan P, Pogue J, Micks M, et al; Heart Outcomes Prevention Evaluation (HOPE) 2 Investigators. Homocysteine lowering with folic acid and B vitamins in vascular disease. *N Engl J Med*. 2006;354:1567–77. [PubMed ID: 16531613]
4. Dormandy JA, Charbonnel B, Eckland DJ, Erdmann E, Massi-Benedetti M, Moules IK, et al; PROACTIVE INVESTIGATIONS. Secondary prevention of macrovascular events in patients with type 2 diabetes in the PROactive Study (PROspective pioglitAzone Clinical Trial In macroVascular Events): a randomised controlled trial. *Lancet*. 2005;366:1279–89. [PubMed ID: 16214598]

5. Hansson L, Hedner T, Lund-Johansen P, Kjeldsen SE, Lindholm LH, Syvertsen JO, et al. Randomised trial of effects of calcium antagonists compared with diuretics and beta-blockers on cardiovascular morbidity and mortality in hypertension: the Nordic Diltiazem (NORDIL) study. *Lancet*. 2000 Jul 29; 356:359–65. [PubMed ID: 10972367]
6. Pitt B, Poole-Wilson PA, Segal R, Martinez FA, Dickstein K, Camm AJ, et al. Effect of losartan compared with captopril on mortality in patients with symptomatic heart failure: randomised trial—the Losartan Heart Failure Survival Study ELITE II. *Lancet*. 2000 May 6;355:1582–7. [PubMed ID: 10821361]
7. Ekblom T, Dahlöf B, Hansson L, Lindholm LH, Scherstén B, Wester PO. Antihypertensive efficacy and side effects of three beta-blockers and a diuretic in elderly hypertensives: a report from the STOP-Hypertension study. *J Hypertens*. 1992 Dec;10(12):1525–30. [PubMed ID: 1338084].
8. Hansson L, Lindholm LH, Ekblom T, Dahlöf B, Lanke J, Scherstén B, Wester PO, Hedner T, de Faire U for the STOP-Hypertension-2 study group. Randomised trial of old and new anti-hypertensive drugs in elderly patients: cardiovascular mortality and morbidity the Swedish Trial in Old Patients with Hypertension-2 study. *Lancet*. 1999 Nov 20;354(9192):1751–6. [PubMed ID: 10577635]
9. Yusuf S, Sleight P, Pogue J, Bosch J, Davies R, Dagenais G. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. The Heart Outcomes Prevention Evaluation Study Investigators. *N Engl J Med*. 2000;342:145–53. [PubMed ID: 10639539]
10. A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). CAPRIE Steering Committee. *Lancet*. 1996 Nov 16;348(9038):1329–39. [PubMed ID: 8918275]
11. Black DM, Cummings SR, Karpf DB, Cauley JA, Thompson DE, Nevitt MC, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. *Lancet*. 1996 Dec 7;348(9041):1535–41. [PubMed ID: 8950879]
12. Bennett CL, Silver SM, Djulbegovic B, Samaras AT, Blau CA, Gleason KJ, et al. Venous thromboembolism and mortality associated with recombinant erythropoietin and darbepoetin administration for the treatment of cancer-associated anemia. *JAMA*. 2008;299:914–24. [PubMed ID: 18314434]
13. Melander H, Ahlqvist-Rastad J, Meijer G, Beer-mann B. Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. *BMJ*. 2003;326:1171–3. [PubMed ID: 12775615]
14. ISIS-4: a randomised factorial trial assessing early oral captopril, oral mononitrate, and intravenous magnesium sulphate in 58,050 patients with suspected acute myocardial infarction. ISIS-4 (Fourth International Study of Infarct Survival) Collaborative Group. *Lancet*. 1995;345:669–85. [PubMed ID: 7661937]
15. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17,187 cases of suspected acute myocardial infarction: ISIS-2. ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. *Lancet*. 1988;2:349–60. [PubMed ID: 2899772]
16. Maggioni AP, Franzosi MG, Fresco C, Turazza F, Yognoni G. GISSI trials in acute myocardial infarction. Rationale, design, and results. *Chest*. 1990;97(4 Suppl):146S–150S. [PubMed ID: 18436948]
17. Guyatt G, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, Schünemann HJ; GRADE working group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ*. 2008;336:924–6. [PubMed ID: 18436948]