

1.8 Metod

Bakgrund

För varje SBU-rapport utgör litteraturgranskningen det viktigaste hjälpmedlet för att sammanfatta kunskapsläget. Avsikten med granskningen är att sammanställa data baserade på kritisk läsning och granskning av vetenskapliga studier. Kapitel 2–6 bygger huvudsakligen på systematiska och kritiskt sammanställda litteraturgenomgångar och redovisar punktvis slutsatser graderade efter vetenskaplig styrka. Kapitel 1 samt Kapitel 2.1, 3.1, 3.3, 3.5 och 3.10 utgör översiktliga bakgrundsbeskrivningar av olika aspekter på problemområdet och presenterar inga slutsatser.

Val av effektmått

Att endast mäta så kallade surrogatmått är att värdera en begränsad och kliniskt irrelevant aspekt på sjukdomen [2]. Kliniskt viktiga mått på behandlingens nytta, patientrelaterade effektmått, kan vara sjukdomsrelaterad livskvalitet, dödlighet eller symtom av sjukdomen eller dess behandling. Inom området venös tromboembolism, och speciellt dess prevention, har stor del av den publicerade forskningen använt påvisandet av asymtomatiska trombosor; t ex med flebografi som effektmått. Detta har diskuterats i detalj i föregående Kapitel, 1.7. De effektmått gruppen fokuserat på sammanfattas i Faktaruta 1.

Inklusionskriterier

Kraven för inklusion i rapporten var att studier avseende profylax eller behandling skulle:

- ha en randomiserad och kontrollerad design, eller
- utgöra kvalitetsgranskade systematiska litteraturöversikter med metaanalyser, där sådana är lämpliga

- rapportera något av de fyra primära effektmåtten (Faktaruta 1), eller om sådana saknades
- rapportera det sekundära effektmåttet asymtomatisk tromboembolism

Studier rörande terapier eller diagnostiska metoder som i dag inte längre används i Sverige, t ex impedanspletysmografi eller termografi, har endast i undantagsfall inkluderats i rapporten.

För diagnostiska studier krävde gruppen en oberoende och blindad jämförelse med referensmetoden; flebografi, lungarteriografi eller lungskintigrafi tolkad enligt PIOPED, eller en långsiktig klinisk uppföljning av patienter med negativt testutfall som inte fick behandling (se nedan och Kapitel 3.1).

Litteratursökning

Sökning utfördes av SBU:s informatiker i Medline, i vissa fall kompletterat med Cochranebiblioteket, från 1966 till årsskiftet 2000–2001. Därefter har viss komplettering skett vilket framgår av de enskilda avsnitten. Vi har återkommande sökt i Cochranebiblioteket för att identifiera systematiska översikter utförda av Cochranesamarbetet samt översikter som kvalitetsgranskats av Center for Review and Dissemination vid Universitet i York (databasen DARE). Vid behov har vi också sökt i Cinahl, Embase, Swemed och Spriline. Studier på engelska, skandinaviska språk, tyska och franska har inkluderats.

Den primära söktermen var "thrombophlebitis or thromboembolism or pulmonary embolism" (MeSH) kompletterat med "subheadings dt, th, su, rh och dh". Specifika metoder har definierats med egna MeSH-termer i de enskilda avsnitten. "Publication type" har angetts till RCT (randomiserade och kontrollerade kliniska studier), CCT (kontrollerade kliniska studier) och "meta-analysis". Referenslistor i granskade artiklar har varit en viktig källa liksom kongressrapporter och gruppmedlemmarnas egna kontakter.

Företag med heparin eller LWMH-preparat fick en förfrågan om eventuell opublicerad dokumentation utan att några opublicerade studier kom till vår kännedom och utfallet förändrade inte våra resultat eller tolkningar [4].

Metod för granskning

Två personer granskade var för sig abstraktlistorna. Alla artiklar som minst en av granskarna bedömde vara av möjlig relevans lästes därefter i sin helhet. Utifrån denna andra omgång sorterade på nytt de två granskarna bort artiklar därför att de:

- inte berörde ämnet
- inte svarade på våra frågeställningar, dvs inte redovisade något av våra angivna effektmått
- profylax, diagnostik eller behandling av venös tromboembolism inte var huvudobjekt
- var skrivna på annat språk än svenska, danska, norska, engelska, tyska och franska.

Utfallet av dessa procedurer registrerades liksom utfallet av dubbelbedömningen. De preliminärt godkända studierna gick slutligen till en strukturerad granskning och dataextraktion enligt nedanstående.

Kvalitetsvärdering av de vetenskapliga studierna

Väsentligen baseras kunskapssammanställningen på randomiserade, kontrollerade studier, systematiska översikter med metaanalyser eller diagnostiska studier med en oberoende jämförelse med en referensmetod eller med en klinisk långtidsuppföljning. Endast i de fall vi inte kunnat identifiera randomiserade och kontrollerade studier har vi inkluderat t ex kontrollerade studier men utan randomisering eller fallserier. Bevisvärdet av sådana är begränsat och de redovisas därför endast översiktligt. Vi har genomgående använt dubbelgranskning av studier och gruppens

medlemmar har inte ensamma fått granska studier där de själva deltagit. De mallar som använts som underlag för granskning och dataextraktion visas i Appendix.

Granskningen redovisas i tre delar [5,6]:

- beskrivning av studiens design och hur diagnostik och intervention genomförts samt i vilka patientgrupper och under vilka omständigheter
- värdering av studiens interna validitet (tillförlitlighet i design, sannolikhet för avsaknad av systematiska fel, bias)
- extraktion av data avseende effektens storlek, precision och riktning.

Rapporten ger en utförlig beskrivning av patientmaterialets sammansättning eftersom detta ger läsaren möjlighet att bedöma om patienterna som studerats, stämmer överens med den kliniska vardag läsaren befinner sig i, även benämnt extern validitet. I studier som använt omfattande exklusionskriterier kan man inte extrapolera, eller generalisera, mått såsom absolut riskreduktion (ARR) eller ”number needed to treat” (NNT) till en klinisk vardag. Om studiepopulationen dessutom selekterats så att en mycket speciell biologisk domän studerats, kan kanske inte heller relativa riskmått extrapoleras.

Det väsentligaste momentet i granskningen är värdering av risken för systematiska fel (bias) genom brister i design och genomförande av studien. Många olika sorters poängsystem har förespråkats för att genom checklistor och poäng kunna skilja ut studier med låg risk för bias från de med hög risk. En nyligen publicerad sammanställning av 25 sådana skalor visar att man genom val av skala kan få vilket svar man önskar när man gör en metaanalys med åtföljande kvalitetsklassning [1,7]. Vi har därför avstått från att använda poängsättning av studier.

Det finns tre grundläggande element som har ett empiriskt stöd för att minska risken för systematiska fel och granskningen fokuserar därför på dessa:

- att deltagarna randomiserats till den aktiva behandlingen (interventionsgrupp, i tabeller förkortat I) eller kontrollgrupp (i tabellerna förkortat K, vilket kan vara ett annat aktivt preparat eller aktiv behandling alternativt placeboberedning)
- att uppföljning av de som ursprungligen randomiserades är så komplett som möjligt och att de som fallit bort har inkluderats i sin ursprungliga grupp, analys enligt ”intention to treat”
- att patienter, forskare och den som mäter behandlingseffekten gör detta på samma sätt i de båda studiearmarna. Detta förutsätter ofta att undersökarna är blindade (dvs är okunniga om) avseende vilken behandling patienten randomiserats till.

Brott mot dessa principer, speciellt de första två, är oftast förenat med en felaktig uppskattning av storleken på behandlingseffekten. Hur långt blindningsprocedurer ska drivas, och vem som behöver vara blindad, är beroende av vilken typ av profylax eller behandling som studeras. I de flesta studier av venös tromboembolism bör såväl behandlande läkare och sjuksköterska som patient vara blindade. Den som utvärderar behandlingseffekten, t ex genom klinisk undersökning, uppskattning av blödning eller granskning av flebografibilder, bör också vara okunnig om vilken grupp patienten tillhör. I vissa sammanhang är det dock omöjligt att personal och patienter är okunniga om vilken behandling som ges (t ex vid användandet av kompressionsstrumpor).

För diagnostiska studier krävde gruppen en oberoende, blindad jämförelse med en referensmetod (”gold standard”) samt att patientpopulationen inkluderade ett rimligt spektrum av patienter motsvarande de på vilka testen ska appliceras i klinisk vardag. Även så kallade managementstudier, med långsiktig klinisk uppföljning av patienter med negativt testutfall som inte fick behandling samt objektiv diagnostik av nya episoder, inkluderades. Detta diskuteras vidare i Kapitel 3.1. Konsekutiv inklusion av patienter var önskvärd då det kan vara en indikation på god generaliserbarhet under de omständigheter som undersökningen genomförts. Diagnostiska test har, om möjligt, analyserats med avseende på

användbarhet i journalsituationer samt variabilitet mellan olika undersökare. Granskningsmallen för diagnostiska studier återges i Appendix.

Systematiska litteraturoversikter, med eller utan kvantitativ metod (metaanalys), inkluderades om de hade använt adekvata inklusionskriterier, täckt de flesta relevanta studierna samt värderat tillförlitlighet och reproducerbarhet. En väl genomförd metaanalys bör värdera grad av homogenitet i resultaten och utforska och tolka en eventuell heterogenitet.

Redovisning av effekter

Slutligen extraherade vi data avseende behandlingseffektens storlek och riktning. Vi har bemödat oss om att redovisa absoluta tal samt kalkylera:

- absolut riskreduktion (ARR)
- relativ riskreduktion (RRR)
- number needed to treat (NNT).

För ogynnsamma utfall (t ex blödningar) redovisas:

- absolut riskökning (ARÖ)
- relativ riskökning (RRÖ)
- number needed to harm (NNH).

I tabeller anges påvisade effekter i första hand som risk i de enskilda grupperna samt som absolut riskreduktion (dvs skillnaden mellan grupperna), men p-värden anges oftast inte. I stället anger vi, där så har varit möjligt, de 95-procentiga konfidensintervallen (KI) för skillnaden (ARR eller ARÖ). I rapportens text beskrivs konfidensintervallen som (x till y) och i tabeller som (x ; y).

Vi vill i detta sammanhang prioritera att visa klinisk relevans framför statistisk signifikans, som dock indirekt framgår av konfidensintervallen. Vi kan därmed undvika att dikotomisera resultaten i signifikant/inte signifikant och lämna till läsaren att värdera om det uppgivna konfidensintervallet sannolikt är förenligt med en rimlig nytta eller risk. Det ger

också möjlighet att värdera även traditionellt ”icke-signifikanta” resultat. Vi anger inte relativa riskreduktioner eller NNT i de fall konfidensintervallen omfattar värdet för ingen effekt (dvs noll för ARR och 1 för OR eller RR) eller p-värdet är större än 0,05.

Vi har genomfört några egna metaanalyser med hjälp av programvaran Comprehensive Meta-analysis (Biostat, Englewood, NJ, USA) och för diagnostiska studier använde vi ett program konstruerat av Josef Lau vid Cochranesamarbetet.

Att syntetisera och styrkegradera en slutsats

Efter att ha granskat de studier som är relevanta och uppfyller basala kvalitetskrav har vi vägt samman studierna och syntetiserat en slutsats. I denna process måste såväl studiekvalitet som samstämmighet mellan olika studier (konkordans) vägas in. Slutsatser har ett starkt empiriskt stöd om de baseras på många bra och stora studier med likartade resultat, avseende storlek och riktning av effekten. Få och små studier med kvalitetsbrister och heterogena resultat (diskordans) kan aldrig leda till annat än mycket svaga, om än några, enhetliga slutsatser.

SBU:s projektgrupper redovisar slutsatserna av varje avsnitt med en fyrgradig gradering av det empiriska underlaget från kliniska studier. Metoden har utvecklats i samarbete med Läkemiddelsverket och Socialstyrelsen. Denna process bygger på att projektgruppen först beslutar sig för en klassning av bevisvärdet av studier inom det aktuella området och för de frågeställningar som gruppen formulerat. Gruppens klassificering av de ingående studiernas bevisvärde framgår av Faktaruta 2–4.

I värderingen av det empiriska underlaget har vi använt begreppet bevisvärde, eftersom detta har blivit allmänt accepterat av SBU, Socialstyrelsen och Läkemiddelsverket. Det är ur många synpunkter praktiskt med en samlande term, men med nödvändighet inträder en del svårigheter när slutprodukten av en så komplex process som kunskapsutveckling ska beskrivas med ett ord. Svårigheterna blir mest markanta när vi rör oss med relativt eller helt nya rön i kunskapsfronten.

Den förmodligen största risken för missförstånd ligger i att ordet bevis för många är knutet till ”bevis för att något är sant”. Emellertid finns inget konsensus för hur ett sanningskriterium ser ut i vetenskapen. Vetenskapen prövar hela tiden sina gamla sanningar och söker nya, bättre förklaringar och utvecklingen visar just att ”sanningar” och därmed hela deras bevisgrund hela tiden måste omprövas för att framsteg ska kunna göras. ”Bevisvärde” i den bemärkelse vi använder det i kunskapsöversikten får därför mer uppfattas som en gradering av hur väl en teori, behandling eller annan intervention kritiskt prövats i välgjorda studier och klarat sig hittills – men med de viktiga tilläggen att bevisen kan falla imorgon, att ingen vet sanningen och att all kunskap är provisorisk.

För varje enskild fråga som vår grupp har besvarat (t ex reducerar profylax med lågmolekylärt heparin dödligheten hos patienter med stroke?) har vi vägt samman de tillgängliga studierna. För många av frågorna fanns bara ett fåtal, eller inga, sådana studier. Från sammanställningen har slutsatsen underbyggts med olika styrka.

I Faktaruta 5 redovisas det principiella schemat för en sådan gradering. Om underlaget är mycket bristfälligt eller helt saknas (mindre än Grad 3) har vi inte dragit några slutsatser. Viktigt är dock att en sådan gång hålla isär de två begreppen ”avsaknad av bra empiriskt underlag”, alltid grad fyra, från ”belägg för att effekt saknas”, en slutsats som kan graderas från ett till tre. Enstaka frågor är studerade i flera studier men med så motsägelsefulla resultat att vi valt att klassa vårt utlåtande som ”Resultaten är motstridiga för att värdera ... (3)”.

Många komponenter blir nödvändigtvis i viss mån subjektiva – vad är en stor respektive liten studie och hur mycket diskordans kan man acceptera. Genom gemensamma diskussioner i gruppen har vi uppnått en rimlig samsyn på tolkningsproblemen. Ju tunnare det empiriska underlaget är desto större utrymme finns det för olika tolkningar och man närmar sig snarare ett ”expertutlåtande” än en empiriskt välgrundad slutsats. Vi har syftat till att tydligt ange var vi anser att kunskaperna är otillräckliga för riktigt pålitliga slutsatser.

Styrkegradering av slutsatser får inte tolkas som absoluta sanningar eller storheter utan är att betrakta som en hjälp för läsaren och de som ska använda rapporten för att fatta beslut om riktlinjer för en klinisk vardag. Självfallet kan slutsatser som baseras på Grad 1 leda till mer konkreta anvisningar om patientbehandlingar än Grad 3.

Metod för de hälsoekonomiska analyserna

Urval

Från databasen Medline söktes artiklar publicerade åren 1996–2001 med sökorden ”venous embolism, pulmonary embolism, deep venous thrombosis” samt ”cost(s), health economics, cost-effectiveness analysis”. Ytterligare sökningar gjordes i Embase och den brittiska databasen NHS Economic Evaluation Database, vilket gav en handfull ytterligare artiklar. Databaserna överlappar varandra till stor del. Efter en första genomgång där kommentarer och brev uteslöts återstod 74 originalartiklar. Artiklarna grupperades efter om de avsåg diagnostik (19), behandling (25) eller profylax (30). Samtliga artiklar granskades och sammanfattades med uppgifter om typ av hälsoekonomisk analys, typ av studie, plats och år för studien, antal patienter, uppföljningstid, effekter, kostnader, resultat, känslighetsanalys, eventuellt finansiellt stöd samt kommentar.

Kriterier

Kriterier för bedömning av hälsoekonomiska studier av profylax, behandling och diagnostik överensstämmer med de som använts för granskningen av metodernas effekter som redovisas i inledningen till detta kapitel.

Innan de hälsoekonomiska studierna har granskats har därför en avstämning gjorts mot de källor som använts och som även kan ha bedömts i respektive huvudavsnitt. Om originalstudien(-erna) ansetts ha ett ringa bevisvärde eller metoden inte används i Sverige har någon hälsoekonomisk granskning inte gjorts. Efter avstämning mot de kliniska bedömningarna samt en första granskning återstod 50 av 74 artiklar.

Hälsoekonomiska analyser jämför alternativ. I samtliga dessa ingår kostnader som beräknas genom att resursförbrukningen identifieras, kvantifieras och kostnadsberäknas. Kostnader kan begränsas till direkta sjukvårdskostnader eller även avse indirekta kostnader som uppstår pga sjuklighet eller tidig dödlighet.

Det finns fyra typer av analyser. De skiljer sig främst i hur effekterna beaktas. I den enklaste typen, kostnads-minimeringsanalys (CMA), antas att effekterna är lika och de ingår därför inte. I kostnads-effektanalys (CEA) jämförs kostnader med effekter som mätts i naturliga fysiska enheter (t ex antal funna fall). Livskvalitet kan även jämföras och dessa analyser benämns kostnads-nyttoanalyser (CUA). Därvid jämförs kostnader med ett effektmått som både inkluderar kvantitet och kvalitet, exempelvis kvalitetsjusterade levnadsår (QALY). En fjärde typ av analys värderar även effekterna i ekonomiska termer, kostnads-intäktsanalys (CBA).

Hälsoekonomiska analyser kan bygga på data som insamlats direkt i samband med kliniska studier. Fördelarna är flera, t ex att resurs-utnyttjandet avser just de personer som även effekterna redovisas för. Är studien en RCT ökar generaliserbarheten. En nackdel är att studierna ofta är för små för att medge statistisk analys av skillnader i kostnader. Uppföljningen omfattar också ofta kort tid, veckor eller några månader.

Många hälsoekonomiska studier är genomförda med hjälp av modeller. Dessa byggs upp för att simulera hur effekter och kostnader förändras över tiden. Sannolikheten för olika utfall tas från redan publicerade studier. Kostnader hämtas ofta från helt andra studier. För att efterlikna den kliniska verkligheten kan ett stort antal hypotetiska patienter analyseras med avseende på kostnader och utfall. Fördelarna är att redan tillgänglig kunskap kan utnyttjas. Nackdelen är att såväl effekter som kostnader kan ha helt olika källor och ofta analyseras bara ett enda typfall.

Många studier anger att en viss metod är kostnadseffektiv. Begreppet kostnadseffektivitet är i sig inte särskilt informativt då det innehåller två komponenter, skillnader i effekter respektive skillnader i kostnader

för en alternativ metod jämfört med utgångsläget. Därmed kan fyra olika kombinationer av ändrade kostnader och ändrade effekter uppstå. Oftast medför en ny metod något ökad effekt men till priset av även ökade kostnader. Det finns inga fastställda eller överenskomna gränser när en undersökning eller behandling kan anses vara kostnadseffektiv. I stället konstateras i efterhand att en del metoder som införts visat sig ha en kostnad per t ex extra år av vunnet liv uppemot 200 000–400 000 SEK.

Bevisvärde

För att bedöma de hälsoekonomiska studiernas bevisvärde utnyttjades en checklista (Faktaruta 6) [3]. Den omfattar tio frågor som belyser olika aspekter av hälsoekonomiska studier. Det finns en rangordning av dessa frågor.

Den viktigaste frågan är om den hälsoekonomiska studien ställt en för svensk sjukvård relevant fråga som även kan besvaras med hjälp av en hälsoekonomisk studie.

Hälsoekonomi avser underlätta beslut och därför krävs även att studien har en fullständig redovisning av olika alternativ. För alternativen ska både effekter och resursåtgång identifieras, kvantifieras och värderas.

En tredje grupp av frågor avser resultaten och deras stabilitet. Alla analyser bygger på vissa antaganden och det är viktigt att det redovisas hur resultaten förändras om antagandena ändras, hur de förändras på marginalen samt över tid. Önskvärt är även en statistisk analys av skillnader i kostnader för olika alternativ att nå vissa effekter.

En fjärde typ av frågor avser hur de redovisade resultaten kan omsättas i praktiken som underlag för de som ska fatta beslut. Studiedesign och resultat är mycket starkt kopplade till de sjukvårdssystem och den epok där de genomförts. Därtill har studierna ofta genomförts under optimala villkor. Detta innebär att studierna oftast visar mest gynnsamma utfall i just den miljö de genomförts. Generaliserbarheten av resultaten för en svensk sjukvårdsmiljö är därmed ofta lägre.

De hälsoekonomiska studiernas bevisvärde har bedömts enligt ovanstående prioritet för de tio frågorna. Studiernas bevisvärde har därefter grupperats enligt kriterierna i Faktaruta 7.

I respektive avsnitt om hälsoekonomi vid profylax (2.7), diagnostik (3.11) och behandling (4.8) har studierna grupperats, sammanfattats och bedömts. Därefter har bedömningarna granskats, stämts av mot de epidemiologiska avsnitten och deras vetenskapliga styrka graderats. Detta har gjorts enligt en skala med fyra grupper, analog med den som använts för den kliniska bedömningen (Faktaruta 5).

Faktaruta 1

De primära effektmåtten som ligger till grund för urval av studier och slutsatser vid profylax och behandling av venös tromboembolism:

- dödlighet, total respektive till följd av lungemboli eller dödliga blödningar
- morbiditet i form av symtomgivande DVT, LE eller posttrombotiskt syndrom (PTS) verifierade med objektiva metoder
- morbiditet i form av icke-dödliga blödningar (allvarliga/icke allvarliga)
- klinisk långtidsuppföljning av patienter (managementstudier)

Sekundära effektmått:

- asymtomatisk DVT och LE.

Faktaruta 2

Underlag till klassning av bevisvärdet av studier rörande profylax eller behandling.

Högt bevisvärde

- Randomiserade studier med blind design och utvärdering samt rimligt statistiskt underlag. Väl genomförda managementstudier med objektiv validering av kliniska händelser under uppföljningstiden.

Medelhögt bevisvärde

- Öppna randomiserade studier eller där studiestorleken inte baserats på ett rimligt statistiskt underlag.
- Randomiserade studier med ”smärre” brister i randomisering eller uppföljning.
- Välgjorda icke-randomiserade kohortstudier med klart definierade inklusionskriterier, exponeringar och effektmått och samtidiga jämförelsegrupper.
- Välgjorda fall–kontrollstudier med klart definierade inklusionskriterier, exponeringar och effektmått.

Lågt bevisvärde

- Randomiserade studier med avgörande brister i design och genomförande.
- Alla övriga icke-randomiserade studier.

Faktaruta 3

Underlag till klassning av bevisvärdet av systematiska översikter.

Högt bevisvärde

- Översikter som redovisar adekvata inklusionskriterier, värderat studiernas tillförlitlighet och homogenitet samt är reproducerbart tolkade. Litteratursökning som förutom publicerade studier inkluderar minst en av abstrakten, kontakt med författare eller företag. Eventuell metaanalys utförd med analys av heterogenitet samt sensitivitetsanalys.

Medelhögt bevisvärde

- Översikter med "smärre" defekter i något av: inklusionskriterier, värdering av studiernas tillförlitlighet och homogenitet samt reproducerbarhet i tolkning. Litteratursökning inkluderar enbart publicerade studier. Metaanalys med brister i statistik.

Lågt bevisvärde

- Översikter som inte är systematiska i sin design eller som felaktigt använder metaanalys när så inte är befogat.

Faktaruta 4

Underlag till klassning av bevisvärdet av diagnostiska studier.

Högt bevisvärde

- Oberoende, blindad jämförelse med en referensmetod i en patientpopulation som inkluderar ett rimligt spektrum av konsekutiva patienter. Referensmetod ska ha använts till alla patienter eller om referensmetoden inte använts till alla patienter ska uppföljning skett på ett adekvat sätt (managementstudier). Data redovisade i fyrfältsform eller som sensitivitet och specificitet och adekvat metodbeskrivning.

Medelhögt bevisvärde

- Oberoende, blindad jämförelse med en referensmetod i en patientpopulation som inte är konsekutivt inkluderade.

Lågt bevisvärde

- Övriga designern.

Faktaruta 5

Styrkegradering av slutsatser

Grad 1 Starkt vetenskapligt underlag

Minst två oberoende studier med högt bevisvärde eller en systematisk översikt av högt bevisvärde.

Följande undantag från huvudregeln kan också motivera Grad 1:

Enstaka mycket stor, väl designad och genomförd RCT utförd på ett stort antal centra.

Grad 2 Måttligt starkt vetenskapligt underlag

En studie med högt bevisvärde plus minst två med medelhögt bevisvärde och entydiga resultat (huvudregel).

Följande undantag kan förekomma:

Väl upplagda kohortstudier eller fall–kontrollstudier med entydiga resultat.

Ett flertal RCT med medelhögt och lågt bevisvärde med entydigt resultat.

Minst en diagnostisk studie med högt bevisvärde och flera med medelhögt bevisvärde.

En stor diagnostisk studie med högt bevisvärde.

En systematisk översikt med medelhögt bevisvärde med stöd från oberoende randomiserade studier.

Grad 3 Begränsat vetenskapligt underlag

Minst två studier med medelhögt bevisvärde och entydigt resultat.

Följande undantag kan förekomma:

Väl upplagda kohortstudier eller fall–kontrollstudier med divergerande resultat.

Ett flertal RCT med medelhögt eller lågt bevisvärde.

En diagnostisk studie med högt bevisvärde men begränsad statistisk styrka.

En systematisk översikt med medelhögt bevisvärde utan stöd från oberoende studier.

Grad 4 Otillräckligt vetenskapligt underlag

Enbart studier av lågt bevisvärde eller avsaknad av studier.

Expertutlåtanden eller konsensusutlåtande utan explicit systematisk genomgång.

OBS Vid heterogena resultat mellan studier som inte kan förklaras utifrån patientmaterial eller design reduceras styrkegraden åtminstone en nivå.

Faktaruta 6

Checklista för granskning av hälsoekonomiska studier [3]

- 1. Föreligger en relevant frågeställning som också kan besvaras?**
 - a) Studeras både kostnader och effekter?
 - b) Studeras alternativ(a) behandling(ar)?
 - c) Anges i studien för vilken beslutssituation som studien har genomförts?
- 2. Föreligger en utförlig beskrivning av behandlingsalternativ?**
 - a) Har något viktigt alternativ inte tagits med?
 - b) Har noll-alternativet beaktats?
- 3. Hur har behandlingsalternativens effekter dokumenterats?**
 - a) Randomiserad, kontrollerad studie?
 - b) Kontrollerad studie?
 - c) Annan studie?
- 4. Har alla väsentliga och relevanta kostnader och effekter för samtliga alternativ identifierats?**
 - a) Har alla relevanta perspektiv beaktats?
 - b) Har såväl rörliga som kapitalkostnader inkluderats?
- 5. Har kostnader och effekter kvantifierats och värderats på ett relevant sätt i rimliga fysiska enheter?**
- 6. Har kostnader och effekter värderats på ett trovärdigt sätt?**
- 7. Har kostnader och effekter diskonterats?**
 - a) Diskonterade till nuvärdet?
 - b) Är diskonteringsräntan rimlig?
- 8. Har en marginalanalys gjorts av kostnader och effekter?**
- 9. Har känslighetsanalys gjorts?**
 - a) Har lämpliga statistiska analyser utförts?
 - b) Är den tillämpade spridningen för variabelvärdena rimlig?
 - c) Är studiens resultat känsliga för varierande variabelvärden?
- 10. Har presentationen och diskussionen av studiens resultat inkluderat alla väsentliga aspekter för beslutsfattarna?**
 - a) Tydligt uttryckta slutsatser?
 - b) Jämförs resultaten med andra studiers resultat?
 - c) Diskuteras om resultaten är generaliserbara?
 - d) Har studien beaktat även andra väsentliga förhållanden med avseende på beslutssituationen (t ex etiska frågeställningar)?
 - e) Diskuteras möjligheter till praktiskt genomförande (t ex i form av en konsekvensanalys)?

Faktaruta 7

Bevisvärde av hälsoekonomiska studier

Högt	Adekvat frågeställning, design, redovisning av alternativ samt hög bedömning av övriga aspekter.
Medelhögt	Adekvat frågeställning, design, redovisning av alternativ samt medelhög bedömning av övriga aspekter.
Lågt	Bristfällig redovisning av frågeställning, design, alternativ och låg bedömning av övriga aspekter.

Referenser

1. Berlin JA, Rennie D. Measuring the quality of trials: the quality of quality scales. *JAMA* 1999;282:1083-5.
2. Bucher HC, Guyatt GH, Cook DJ, et al. 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.
3. Drummond MF, O'Brien B, Stoddard G, Torrance G. Methods for the economic evaluation of health care programmes. Oxford University Press 1998.
4. Eliasson M, Bergqvist D. [Research results should be freely accessible! – Case reports demonstrate obstacles to contact with drug industry]. *Läkartidningen* 2001; 98:3913-6.
5. 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.
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. Juni P, Witschi A, Bloch R, Egger M. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;282:1054-60.