

## De praktijk van systematische reviews. VIII. Zoeken en beoordelen van systematische reviews

W.J.J.ASSENDELFT, R.J.P.M.SCHOLTEN, J.L.HOVING, M.OFFRINGA EN L.M.BOUTER

Een systematische review (SR) berust op drie pijlers: een expliciete zoekstrategie, een beoordeling van de kwaliteit van de erin opgenomen onderzoeken en een transparante presentatie van de resultaten. In het afgelopen decennium is de aard van een overzichtartikel ingrijpend gewijzigd. De plaats van verhalende reviews is ingenomen door SR's. De voordelen van een SR zijn evident: het proces is transparant, volledig en reproduceerbaar.<sup>1</sup> Naast een kwalitatieve samenvatting kan een kwantitatieve samenvatting (statistische pooling of meta-analyse) onderdeel zijn van een SR. Het kwantitatief combineren van data uit verschillende onderzoeken vergroot de onderzoekspopulatie en zo ook de mogelijkheid om tot een nauwkeurige effectschatting te komen en om op statistische gronden subgroepen die gevoeliger zijn voor de interventie te kunnen identificeren.<sup>2</sup> Omdat de SR op valide manier de resultaten uit verschillende onderzoeken samenvat wordt deze, indien het een review van gerandomiseerde klinische trials (RCT's) betreft, in de hiërarchie van klinisch bewijs op de hoogste plaats gezet (<http://cebm.jr2.ox.ac.uk/docs/levels.html>). SR's van niet-gerandomiseerde cohort- en van case-controlonderzoeken bevinden zich op een lager niveau van bewijs. Men dient zich echter te realiseren dat voor veel onderwerpen en interventies RCT's simpelweg ontbreken of onmogelijk zijn.<sup>3</sup> Daarnaast zijn ook SR's gevoelig voor bias, zelfs als deze op RCT's zijn gebaseerd.<sup>4</sup>

Het is voor de clinicus, beleidsmaker en onderzoeker belangrijk SR's te kunnen zoeken en de methodologische kwaliteit ervan te kunnen beoordelen. Om reden van overzichtelijkheid bespreken wij hier alleen de SR van RCT's, al zijn veel van de aanbevelingen ook op SR's van niet-RCT's van toepassing.

### ZOEKEN VAN SYSTEMATISCHE REVIEWS

Er zijn verschillende manieren om naar SR's te zoeken (tabel 1).<sup>5-7</sup> De volgende aspecten zijn bij het zoeken van belang: (a) de beschikbaarheid van bronnen; (b) de mate

### SAMENVATTING

– In een systematische review (SR) wordt de beschikbare informatie uit de oorspronkelijke onderzoeken op een transparante, valide en reproduceerbare wijze gezocht, geselecteerd, geëxtraheerd, beoordeeld en samengevoegd. SR's zijn daarvoor een efficiënte en betrouwbare informatiebron voor de clinicus, beleidsmaker en onderzoeker.

– SR's kunnen in diverse bronnen worden geïdentificeerd. De zoekstrategie wordt bepaald door de mate van volledigheid die wordt nagestreefd, de beschikbaarheid van diverse bronnen en de behoefte aan toegevoegd commentaar op de SR door een referent. Voor het zoeken van SR's in *Medline* bestaat een sensitief zoekfilter.

– In een SR dient te worden uitgegaan van een nauwkeurig geformuleerde klinische vraagstelling, die aan de hand van de gegevens uit de onderzoeken ook daadwerkelijk wordt beantwoord.

– De kwaliteit van SR's kan sterk verschillen en dient zorgvuldig te worden beoordeeld. Bij deze beoordeling dient aandacht te worden besteed aan: vraagstelling, zoekactie, selectie van onderzoeken, kwaliteitsbeoordeling, data-extractie en -presentatie, het (statistisch) combineren van gegevens, statistische en klinische heterogeniteit, resultaten en conclusie. Er zijn meerdere beoordelingslijsten beschikbaar.

– Verschillen in conclusies tussen SR's over hetzelfde onderwerp zijn op systematische manier te analyseren en vaak op te lossen.

waarin volledigheid van de bron (zoveel mogelijk SR's) wordt nagestreefd; (c) de behoefte aan een begeleidend commentaar op of interpretatie van een SR door een referent.

Een zoekactie in een elektronische database, zoals *Medline* of *Embase*, wordt aangeraden indien aan de hand van een klinische vraag zo uitgebreid mogelijk naar SR's moet worden gezocht. Voor het zoeken in een elektronische database is enige ervaring vereist. Voor het identificeren van SR's kan een zoekfilter worden gebruikt (tabel 2).<sup>7</sup> *Medline* is via het internet gratis te raadplegen, voor toegang tot *Embase* moet altijd individueel of collectief (bijvoorbeeld door de bibliotheek) worden betaald. Het tijdschriftenbestand dat door *Embase* wordt gedekt, is voor een substantieel deel complementair aan *Medline*, onder andere doordat *Embase* meer op farmacotherapie is gericht en ook meer Europees georiënteerd is.<sup>8</sup> Afhankelijk van het onderwerp is de mate van overlap in geïdentificeerde onderzoeken 10 tot 75%. Een uitgebreide systematische zoekactie behoort derhalve tenminste beide databases te beslaan.<sup>9</sup>

Academisch Medisch Centrum/Universiteit van Amsterdam, Dutch Cochrane Centre, Postbus 22.660, 1100 DD Amsterdam.

Dr.W.J.J.Assendelft, huisarts-epidemioloog (tevens: Divisie Public Health, afd. Huisartsgeneeskunde); dr.R.J.P.M.Scholten, arts-epidemioloog; dr.M.Offringa, kinderarts-epidemioloog (tevens: Emma Kinderziekenhuis, afd. Kindergeneeskunde).

Vrije Universiteit, Instituut voor Extramuraal Geneeskundig Onderzoek, Amsterdam.

Dr.J.L.Hoving en prof.dr.L.M.Bouter, epidemiologen.

Correspondentieadres: dr.W.J.J.Assendelft (w.j.assendelft@amc.uva.nl).

TABEL 1. Bronnen van systematische reviews (SR's)

bron	beschikbaarheid	volle- ledig- heid*	met com- men- taar van refe- renten	opmerkingen
<i>Cochrane Database of Systematic Reviews</i>	abbonement <i>Cochrane Library</i> ; abstracts gratis via internet ( <a href="http://www.cochrane.nl">www.cochrane.nl</a> ) of via <i>Medline</i>	++	nee	hoge kwaliteit van de SR's in deze database is aangetoond <sup>12</sup>
<i>Database of Abstracts of Reviews of Effectiveness</i>	abbonement <i>Cochrane Library</i> ; gratis via internet: <a href="http://nhscrd.york.ac.uk/darehp.htm">http://nhscrd.york.ac.uk/darehp.htm</a>	++	ja	samenvattingen van hoogwaardige reviews met commentaar
elektronische database (bijvoorbeeld <i>Medline</i> of <i>Embase</i> )	<i>Medline</i> onder andere gratis via internet ( <i>PubMed</i> ): <a href="http://www.ncbi.nlm.nih.gov/entrez/query.fcgi">www.ncbi.nlm.nih.gov/entrez/query.fcgi</a> voor toegang tot <i>Embase</i> moet altijd betaald worden	+++	nee	zoekt in geïndexeerde tijdschriften; veelal een samenvatting van de SR beschikbaar; gebruik methodologisch filter voor SR's (zie tabel 2)
Sum Search	gratis via internet: <a href="http://SUMSearch.UTHSCSA.edu/searchform4.htm">http://SUMSearch.UTHSCSA.edu/searchform4.htm</a>	+++	nee	zoekt met ingebouwde filter via internet in <i>Medline</i> en de <i>Cochrane Library</i>
<i>Clinical evidence</i>	abbonement	++	ja	leerboek; aanbeveling wordt gebaseerd op goede SR indien beschikbaar; tweemaal per jaar volledig gereviseerd; uitgangspunt is een klinische vraag
<i>Best Evidence</i>	abbonement	++	ja	tijdschrift; fusie van de tijdschriften <i>ACP journal club</i> en <i>Evidence based medicine</i> ; geen strikte beperking tot SR's, toenemend aantal besproken SR's; uitgangspunt is een verschenen artikel (SR)
tijdschriften	abbonement; sommige tijdschriften full-text op internet (overzicht bijvoorbeeld via <a href="http://www.freemedicaljournals.com">http://www.freemedicaljournals.com</a> )	+	†	bij een klinische vraag is het beter om eerst in <i>Medline</i> of <i>Smart Search</i> te zoeken om vervolgens de betreffende SR's op te zoeken; <i>Huisarts en Wetenschap</i> heeft een speciale Cochrane-rubriek waarin Cochrane-SR's voor de huisarts worden becommentarieerd <sup>11</sup>

\*+++ = zoekt naar zoveel mogelijk SR's; ++ = beperkt aantal SR's; + = slechts SR's verschenen in betreffende tijdschrift.  
†Merendeel bevat geen bijgevoegd referentiencommentaar; sommige tijdschriften wel, bijvoorbeeld *British Medical Journal* en *Huisarts en Wetenschap*.

Ook mag een zoekactie in de *Cochrane Database of Systematic Reviews* niet ontbreken.<sup>1 10</sup> Bij Cochrane-reviews is er een aantal kwaliteitswaarborgen:<sup>11</sup> een protocol dat voor het uitvoeren van de eigenlijke review wordt onderworpen aan beoordeling binnen de eigen reviewgroep (Collaborative Review Group), zorg vanuit diverse subgroepen binnen de Cochrane Collaboration om te waarborgen dat de review aan kwaliteitseisen van diverse eindgebruikers voldoet, regelmatige actualisering en een commentaarpagina met opmerkingen of kritiek bij de review zelf. Cochrane-reviews worden momenteel gezien als de beste samenvatting van het beschikbare klinische bewijs ('best evidence available'). De betere methodologische kwaliteit ten opzichte van niet-Cochrane-reviews is aangetoond.<sup>12</sup> Sinds kort zijn Cochrane-reviews ook in *Medline* geïndexeerd,<sup>13</sup> zodat ze ook bij een zoekactie in *Medline* tevoorschijn komen. Het gratis op internet te gebruiken evidence-based

zoekprogramma Sum Search zoekt naar SR's in *Medline*, maar ook in de *Cochrane Database of Systematic Reviews*. Sum Search vergt minder zoekervaring dan *Medline*, maar is minder geschikt voor meer toegespitste zoekvragen.

Indien er behoefte is aan een toelichting op de SR, kan gezocht worden in de *Database of Abstracts of Reviews of Effectiveness (DARE)*, het leerboek *Clinical evidence* en het tijdschrift *Best Evidence*.<sup>6</sup> In *DARE* worden reviews die aan strenge criteria voldoen, samengevat en van commentaar voorzien. In *Clinical evidence* wordt van groot aantal patiëntgeoriënteerde vragen het beschikbare bewijs (vaak in de vorm van SR's) voor de clinicus geordend en besproken.<sup>14</sup> In *Best Evidence* worden recentelijk verschenen hoogwaardige publicaties (waaronder steeds meer SR's) door experts samengevat en van commentaar voorzien.

Tenslotte verschijnen SR's natuurlijk als afzonderlij-

TABEL 2. Sensitief zoekfilter voor het opsporen van systematische reviews (SR's) in *Medline*

stap	zoekterm	aantal treffers*	opmerkingen
1	meta analysis.pt.	4170	pt ('publication type'): onderzoekstype tw: zoeken met vrije tekstwoorden ('free text words') in de titel en de samenvatting; dit maakt de zoekactie sensitiever, maar minder specifiek; het truncatieteken \$ in vrije tekst betekent dat op alle uitgangen wordt gezocht, dus zowel 'meta-analysis' (enkelvoud) als 'meta-analyses' (meervoud)
2	meta-anal\$.tw.	4890	
3	metaanal\$.tw.	229	
4	metanal\$.tw.	29	
5	quantitativ\$ review\$.tw.	105	
6	quantitativ\$ overview\$.tw.	34	
7	systematic\$ review\$.tw.	1208	
8	systematic\$ overview\$.tw.	95	
9	methodologic\$ review\$.tw.	77	
10	methodologic\$ overview\$.tw.	19	
11	review.pt.	703016	
12	medline.tw.	4230	
13	11 and 12	2539	het totale aantal mogelijke systematische reviews; omdat het een sensitief zoekfilter betreft, zal echter een substantieel aantal niet-SR's worden geïdentificeerd; deze oogst aan mogelijke SR's moet worden gecombineerd met zoektermen die de patiëntencategorie, de interventies en de uitkomst definiëren
14	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 13	10118	

\*Tijdvak: 1966 (start *Medline*)-1999.

ke publicatie in medische tijdschriften. Inhoudsopgaven van beschikbare tijdschriften zijn echter de minst efficiënte manier om naar SR's te zoeken. Bij een gerichte klinische vraag wordt daarom aangeraden om in een elektronische database, bijvoorbeeld *Medline*, te zoeken en vervolgens de SR's in de betreffende tijdschriften op te zoeken.

#### OPBOUW VAN EEN KWALITATIEF GOEDE SYSTEMATISCHE REVIEW

Voor de beoordeling van de methodologische kwaliteit van een SR is het belangrijk alle relevante stappen te begrijpen. De verschillende onderdelen van een SR zijn al eerder in deze reeks in het *Nederlands Tijdschrift voor Geneeskunde* beschreven.<sup>9 15-17</sup> Deze stappen worden voor de lezer van SR's hier samengevat en van praktische adviezen voorzien.

**Vraagstelling.** Een SR dient uit te gaan van een expliciete klinische vraagstelling. De noodzakelijke ingrediënten van een dergelijke vraag zijn: de patiëntencategorie, de interventie, de controle-interventie, de uitkomst en eventueel ook de gewenste follow-upduur. Vaak wordt het Engelse acroniem 'PICO' gebruikt, dat staat voor: 'patients, intervention, control, outcome'. Een voorbeeld van een vraagstelling is: wat is bij patiënten met acute, subacute en chronische tenniselleboog ('patients') de effectiviteit, uitgedrukt in mate van ervaren herstel ('outcome'), van corticosteroïdinjecties ('intervention') ten opzichte van placebo of afwachtend beleid ('control')?<sup>18</sup> Bij deze vraag is gekozen voor globaal herstel als uitkomst en bijvoorbeeld niet voor pijn,

hoewel deze uitkomst ook in een SR had kunnen worden opgenomen.

**Zoekactie.** Op basis van de vraagstelling dient op gestructureerde wijze in de internationale literatuur te zijn gezocht. Een minimumvereiste hiervoor is een zoekactie in de elektronische databases *Medline* en *Embase*, het Cochrane Controlled Trial-register (waarin meer dan 250.000 referenties van trials) en controle van de referenties van de verzamelde artikelen. Optionele aanvullende zoekacties zijn: het raadplegen van (gespecialiseerde) trialregisters en gedrukte indexen, het aanschrijven van experts en fabrikanten en het handmatig doorzoeken van tijdschriften.<sup>9</sup>

**Selectie.** De selectie van de in de SR te includeren artikelen dient bij voorkeur door twee reviewers aan de hand van expliciete in- en exclusiecriteria onafhankelijk van elkaar te zijn geschied.<sup>9</sup> In veel reviews wordt de inclusie beperkt tot publicaties in de talen die de reviewers zelf machtig zijn (meestal het Engels). De kwaliteit van niet-Engelstalige RCT's blijkt echter niet van die van Engelstalige te verschillen, maar de uitkomsten wel: uit verschillende onderzoeken is namelijk gebleken dat de uitkomst van een SR kan afhangen van beperkingen in taal (de zogenoemde 'Tower of Babel bias').<sup>9</sup> SR's met taalrestricties worden derhalve als minder valide beschouwd. Ook het in SR's wel of niet opnemen van ongepubliceerde RCT's kan tot verschil in de uitkomst leiden. Uit onderzoek blijkt namelijk dat in het publicatieproces selectie kan optreden, die leidt tot zogenoemde publicatiebias. De interpretatie van gegevens van ongepubliceerde onderzoeken is niet eenvoudig, met name

door het ontbreken van peer review als kwaliteitscontrole. Vooral nog lijkt de aantrekkelijkste oplossing om wel naar ongepubliceerd materiaal te zoeken en vervolgens een sensitiviteitsanalyse met en zonder de ongepubliceerde gegevens uit te voeren (zie verder). Indien de uitkomsten van deze twee analyses aanzienlijk verschillen, is voorzichtigheid in de interpretatie van de review geboden.

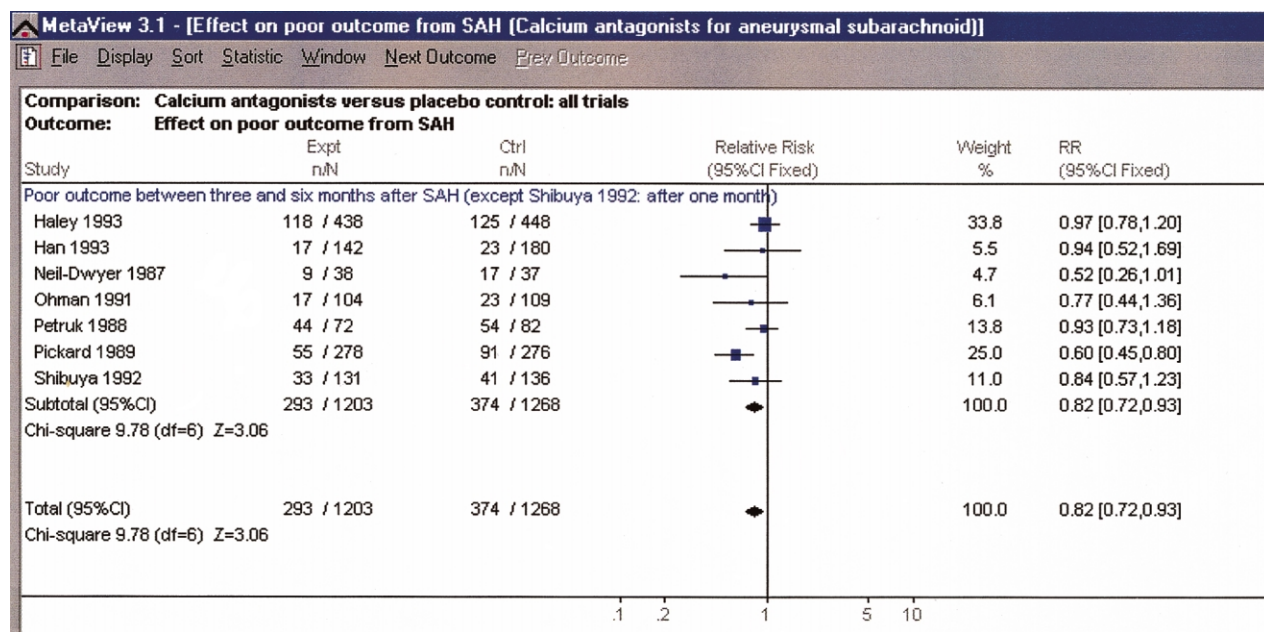
**Kwaliteitsbeoordeling.** De validiteit van de aanbevelingen in een SR hangt sterk af van de kwaliteit van de oorspronkelijke RCT's.<sup>15</sup> Er is een veelheid aan kwaliteitsbeoordelingslijsten beschikbaar.<sup>19</sup> Slechts van een klein aantal interne validiteitsitems is in empirisch onderzoek een vertekende invloed ('bias') aangetoond: blinding van de randomisatie en blinding van patiënten en behandelaars.<sup>15</sup> Een in de SR gebruikte kwaliteitsbeoordelingslijst dient dus tenminste deze items te bevatten. Van de overige items, zoals mate van uitval, is overigens nooit overtuigend aangetoond dat ze geen vertekende invloed hebben. Net als de selectie dient de kwaliteitsbeoordeling in de SR door twee reviewers onafhankelijk van elkaar te zijn uitgevoerd. De resultaten worden vervolgens besproken en er wordt tot consensus gekomen.

**Data-extractie en -presentatie.** In een SR dienen de gegevens die direct bij de vraagstelling aansluiten op een

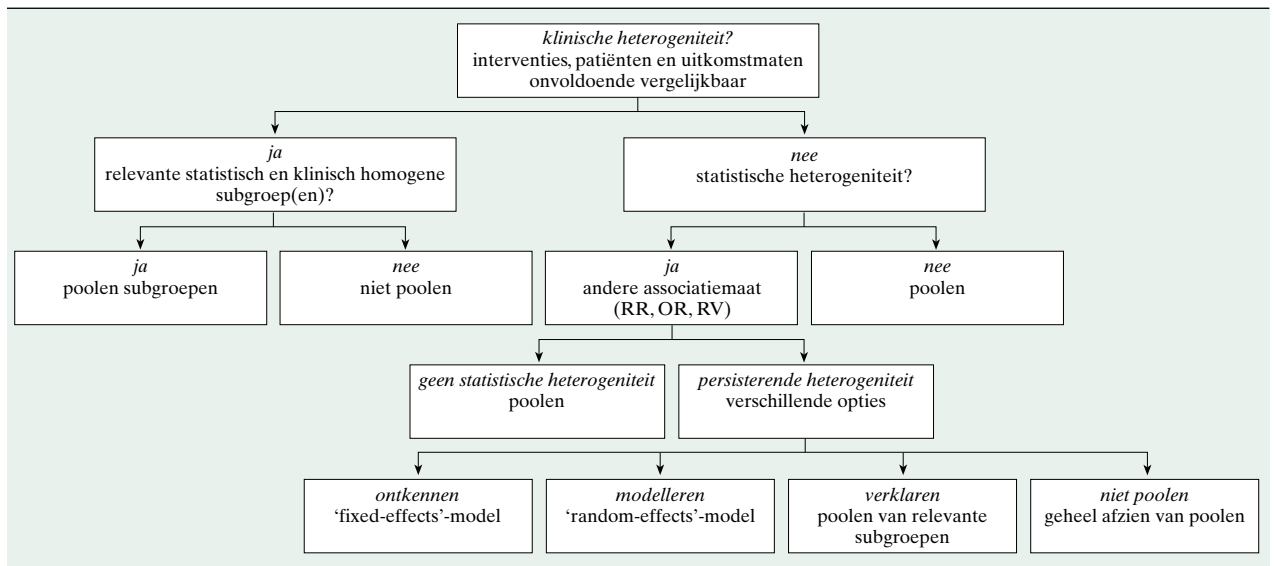
duidelijke wijze te worden gepresenteerd. De gegevens uit de oorspronkelijke RCT's worden in een voor alle RCT's identieke maat uitgedrukt. In een SR wordt dus geen genoegen genomen met de mededeling dat een gevonden verschil in een RCT al dan niet statistisch significant is. Voor dichotome uitkomstmaten (bijvoorbeeld wel/niet genezen) wordt de samenhang uitgedrukt in de oddsratio (OR), het relatieve risico (RR) of het risicoverschil (RV) met het bijbehorende 95%-betrouwbaarheidsinterval (95%-BI), voor continue variabelen (bijvoorbeeld opnameduur in dagen) in een al dan niet gestandaardiseerde verschijscore met 95%-BI. Ook de data-extractie dient bij voorkeur door twee reviewers onafhankelijk van elkaar te geschieden.

**Combineren van resultaten.** In de meta-analyse of de statistische pooling worden de gegevens die voor een systematische review zijn vergaard statistisch samengevoegd om te komen tot een schatting van het in deze onderzoeken bestudeerde effect, in combinatie met de analyse van de invloed van verschillen tussen de onderzoeken op deze schatting.<sup>16</sup> Voor het combineren van effectschattingen ('pooling') zijn diverse statistische technieken beschikbaar.

Met statistische pooling wordt dus een schatting van het effect van een interventie verkregen die gebaseerd is op het grootst mogelijke aantal patiënten, waardoor een



FIGUUR 1. Voorbeeld van een meta-analyse met als onderwerp calciumantagonisten versus placebo voor subarachnoïdale bloeding (SAH), met als uitkomst ongunstige afloop (invaliditeit of overlijden) na 6 maanden, uit de *Cochrane Library of Systematic Reviews*.<sup>20</sup> Ieder horizontaal lijntje geeft de uitkomst weer van een onderzoek, waarbij het lijntje het 95%-betrouwbaarheidsinterval (95%-BI) aangeeft. Het hokje geeft de puntschatter van het relatief risico (RR). Aan ieder onderzoek is een gewicht gegeven dat omgekeerd evenredig is met de variantie: hoe groter het onderzoek, des te kleiner de variantie en derhalve des te groter het gewicht. Het gewicht van het onderzoek staat onder 'weight' en is ook af te lezen aan de grootte van het bij het onderzoek behorende blokje. De ruit is het gecombineerde (statistisch gepoolde) resultaat van de 7 onderzoeken. De verticale lijn betekent een RR van 1. Resultaten die links van de verticale lijn liggen, wijzen op een gunstig effect van de onderzochte therapie ten opzichte van de controlegroep. Calciumantagonisten blijken volgens deze meta-analyse voor subarachnoïdale bloeding effectiever dan placebo: het RR is 0,82 en het 95%-BI sluit de neutrale waarde 1 uit. De kans op een slechte uitkomst (relatieve-risico-reductie) neemt dus met 0,18 (18%) af.



FIGUUR 2. Analyse van heterogeniteit in de resultaten van individuele onderzoeken in een systematische review;<sup>2,17</sup> RR = relatief risico; OR = oddsratio; RV = risicoverschil.

maximale precisie verkregen wordt (smallle betrouwbaarheidsintervallen), zodat ook kleine effecten gedetecteerd kunnen worden. Figuur 1 is een voorbeeld van een meta-analyse naar de effectiviteit van calciumantagonisten versus placebo bij subarachnoidale bloeding, met ongunstige afloop (invaliditeit of overlijden) na 6 maanden als uitkomst.<sup>20</sup>

Pooling is het meest in het oog springende, maar tegelijkertijd ook een veelvuldig verkeerd begrepen en misbruikt onderdeel van de systematische review. Net als bij de analyse van data binnen een RCT zullen statistische softwarepakketten braaf hun werk doen indien de gegevens per onderzoek worden ingevoerd. Voorwaarde voor pooling is echter dat de onderzoeken zowel klinisch als statistisch voldoende gelijk (homogeen) zijn om bij elkaar gevoegd te mogen worden.<sup>17</sup> De inschatting van statistische homogeniteit kan gebeuren met formele statistische toetsing of op het gevoel.<sup>17</sup> De beoordeling van klinische homogeniteit, waarbij de vraag is of de patiënten, interventies en uitkomstmaten voldoende overeenkomen om te mogen worden samengevoegd, is bij uitstek een zaak voor de clinicus. Voor het beoordelen van de klinische relevantie van de patiëntenpopulatie, interventies en effectmaten zijn helaas nog geen goede beoordelingslijsten beschikbaar.

**Heterogeniteit.** Indien de resultaten van de verschillende onderzoeken (sterk) uiteenlopen, spreekt men van heterogeniteit.<sup>17</sup> Heterogeniteit kan berusten op een ongelukkige keuze van het type associatiemaat (de heterogeniteit bestaat bijvoorbeeld wel wanneer de resultaten worden uitgedrukt in het RV, maar niet wanneer uitgedrukt in het RR), verschillen in methodologische kwaliteit tussen onderzoeken of daadwerkelijke verschillen in effectiviteit tussen onderzoeken. Een visuele inspectie van de mate van overlap van de betrouwbaarheidsintervallen van de effectschattingen in de afzonderlijke onderzoeken is de beste manier om heterogeni-

teit tussen onderzoeken te onderzoeken. Door toepassing van het zogenoemde 'random effects'-model voor pooling kan heterogeniteit tussen onderzoeken verdisconteerd worden (figuur 2).<sup>2,17</sup> Het random-effectsmodel levert bij heterogeniteit bredere betrouwbaarheidsintervallen op en is derhalve conservatiever (levert minder snel significante resultaten) dan het 'fixed-effects'-model.

Wanneer er heterogeniteit bestaat tussen onderzoeken in een SR, is verdere analyse van mogelijke bronnen daarvan (subgroepanalyse) geïndiceerd. Dergelijke subgroepanalyses dienen te worden gebaseerd op een aan de SR voorafgaand analyseplan, om 'visexpedities' naar homogene groepen van patiënten of onderzoeken en het daarmee samenhangende risico op fout-positieve bevindingen te voorkomen.<sup>21</sup>

**Sensitiviteitsanalyses.** Na de berekening van een totaaleffect wordt soms nagegaan hoe gevoelig de uitkomst is voor veranderingen in de insluitcriteria voor de meta-analyse of wordt nagegaan of er verschil is tussen RCT's van hoge en van lage kwaliteit.<sup>21</sup> Dit is echter geen 'verplicht' onderdeel van een SR. Als de waarde van de gepoolde schatting hierdoor sterk verandert, dan dient het resultaat met grote voorzichtigheid geïnterpreteerd te worden.

#### BEOORDELINGSLIJST VOOR SYSTEMATISCHE REVIEWS

Het aantal SR's in de internationale literatuur stijgt exponentieel. Echter, de uitkomsten van een SR zijn pas valide en toepasbaar in de praktijk indien de vraagstelling voldoende aansluit bij de vraagstelling in de klinische praktijk, de methoden aansluiten bij de vraagstelling, en als de methodologische kwaliteit van de SR voldoende is. Er bestaat een veelheid aan beoordelingslijsten voor systematische reviews, die soms alle en soms enkele van de voornoemde aspecten van validiteit en toepasbaar-

TABEL 3. Checklist voor de beoordeling van systematische reviews<sup>23 26 27</sup>

item	toelichting
vraagstelling zoekactie	systematische review dient uit te gaan van een expliciete vraagstelling tenminste <i>Medline, Embase, Cochrane Controlled Trial Register</i> en controle van referenties van verzamelde artikelen
selectie	bij voorkeur door 2 reviewers onafhankelijk van elkaar aan de hand van expliciete in- en exclusiecriteria; geen taalrestrictie
kwaliteitsbeoordeling	bij voorkeur door 2 reviewers onafhankelijk van elkaar; expliciet en daardoor reproduceerbaar voor anderen; gevalideerde checklist
data-extractie en -presentatie	resultaten op eenduidige wijze gepresenteerd; per onderzoek wordt per uitkomstmaat een puntschatter van het effect met betrouwbaarheidsinterval gepresenteerd
combineren van resultaten	pooling indien onderzoeken voldoende klinisch en statistisch homogeen zijn; subgroepanalyse gebaseerd op vooraf opgesteld analyseplan
heterogeniteit resultaten conclusie	indien aanwezig: beschreven en adequaat opgelost (zie figuur 2) relevant voor de vraagstelling ondersteund door de resultaten

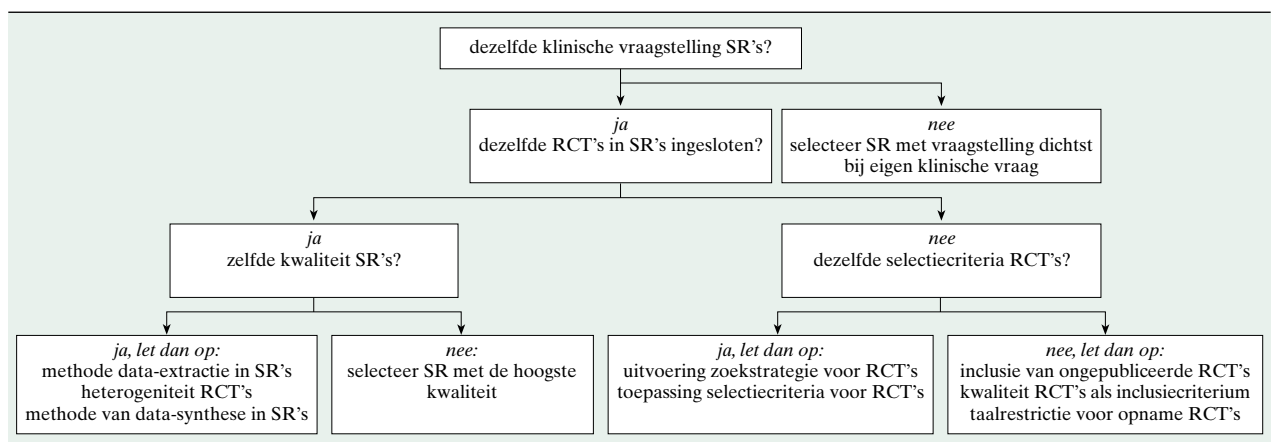
heid bevatten. Shea et al. identificeerden 23 lijsten.<sup>22</sup> De mate van overlap in items is groot. Slechts van een lijst<sup>23</sup> is de validiteit en reproduceerbaarheid onderzocht.<sup>23-25</sup> Onlangs zijn er in de vorm van het QUOROM-statement (van 'quality of reports of meta-analyses') richtlijnen voor systematische reviews vastgesteld en door de meeste toonaangevende tijdschriften onderschreven.<sup>26</sup> (Anders dan in Europa en binnen de Cochrane Collaboration, waar de term 'meta-analyse' gereserveerd is voor de statistische pooling binnen een systematische review, wordt deze aanduiding in de VS vaak als een synoniem voor systematische review gebruikt.) Deze richtlijnen zijn echter vrij omvangrijk en voor de gemiddelde lezer nauwelijks bruikbaar. Derhalve vatten wij de belangrijkste punten samen in tabel 3.<sup>23 26 27</sup>

#### TEGENSTRIJDIGE REVIEWS

Steeds vaker zal de clinicus of de onderzoeker meerdere reviews over hetzelfde onderwerp vinden. De conclusies van verschillende reviews over (schijnbaar) hetzelfde onderwerp zijn soms echter niet eenduidig. De oorzaak van discordantie kan met een algoritme worden ge-

analyseerd (figuur 3).<sup>28</sup> In de regel lukt het aan de hand van dit algoritme om één of meer eenduidige SR's te selecteren en tot een beslissing te komen.

Het is bij discordante SR's over hetzelfde onderwerp allereerst belangrijk te kijken welke SR een vraagstelling heeft die het dichtst bij de klinische vraag (acroniem 'PICO') staat. Indien dezelfde RCT's in de verschillende SR's zijn geïncludeerd, kan men zich richten op de methodologisch beste SR. Indien er geen verschil in kwaliteit is, zal de lezer een nauwkeurige analyse moeten maken van hoe de resultaten zijn geëxtraheerd (bijvoorbeeld: door 2 reviewers onafhankelijk van elkaar), hoe met heterogeniteit is omgegaan en of de uitkomsten op de juiste manier zijn gepoold. Indien niet dezelfde RCT's zijn geïncludeerd, moet men allereerst kijken of bij dezelfde vraagstelling wel dezelfde selectiecriteria voor onderzoeken zijn gehanteerd. Indien dit niet het geval is, zal het wel of niet opnemen van ongepubliceerde onderzoeken het verschil kunnen verklaren. Daarnaast kunnen de SR's onderling verschillen in een eventuele beperking op grond van taal of methodologische kwaliteit van de RCT's. Indien er wat betreft al deze cri-



FIGUUR 3. Algoritme voor het kiezen van de beste passende systematische review(s) (SR('s)) wanneer er over hetzelfde onderwerp SR's met tegenstrijdige conclusies beschikbaar zijn.<sup>27</sup>

teria geen verschil bestaat, kunnen de verschillen in conclusie mogelijk verklaard worden door de verschillen in de praktische uitvoering van het zoeken en het selecteren: vooral de uitvoering hiervan door twee personen, onafhankelijk van elkaar, is belangrijk om het optreden van menselijke fouten te beperken.

#### CONCLUSIE

Steeds meer wordt het klinische handelen gebaseerd op valide, patiëntgebonden onderzoek. De SR verschaft op een transparante en reproduceerbare manier snel inzicht in een grote hoeveelheid informatie. SR's dienen vaak als basis voor richtlijnen, standaarden of protocollen.<sup>29</sup> Indien dergelijke richtsnoeren ontbreken, is de SR een efficiënte en valide manier om inzicht te krijgen in de stand van zaken voor een bepaald onderwerp.

Het is bij een klinische vraagstelling van belang om SR's te selecteren die bij de vraagstelling aansluiten en van een zo hoog mogelijke methodologische kwaliteit zijn. Hiervoor bestaan duidelijke criteria. Gewapend met deze praktische informatie zal de gebruiker snel de voordelen van SR's ontdekken.

---

#### ABSTRACT

*Systematic reviews in practice. VIII. Searching and assessing systematic reviews*

– In a systematic review (SR) the information from the original studies is searched, selected, extracted, appraised and summarised in a transparent, valid and reproducible manner. SRs are therefore an efficient and reliable source of information for the clinician, health policy maker and researcher.

– SRs can be identified in various sources. The search strategy is determined by the desired comprehensiveness, the availability of various sources and the need for additional comments on the SR by a referee. In *Medline* a sensitive filter is available for searching SRs.

– An SR should be based on a carefully formulated clinical question that can be answered on the basis of data from the original studies.

– SRs should be thoroughly assessed as they can differ substantially in terms of quality. In such an appraisal the following aspects should be considered: question, search strategy, selection of studies, quality assessment, data extraction and data presentation, (statistical) summary of the data, statistical and clinical heterogeneity, results and conclusion. Various assessment lists are available.

– Differences in conclusions between SRs on the same subject can be analysed in a systematic manner and can often be resolved.

---

#### LITERATUUR

- 1 Offringa M, Craen AJM de. De praktijk van systematische reviews. I. Inleiding. *Ned Tijdschr Geneesk* 1999;143:653-6.
- 2 Lau J, Ioannidis JPA, Schmid CH. Quantitative synthesis in systematic reviews. *Ann Intern Med* 1997;127:820-6.
- 3 Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* 1996;312:1215-8.
- 4 Ioannidis JPA, Cappelleri JC, Lau J. Issues in comparisons between meta-analyses and large trials. *JAMA* 1998;279:1089-93.
- 5 Sheldon TA, Glanville J. Searching for reliable research evidence need not be difficult [letter]. *BMJ* 1999;319:1435.

- 6 Glanville J, Haines M, Auston I. Finding information on clinical effectiveness. *BMJ* 1998;317:200-3.
- 7 Hunt DL, McKibbin KA. Locating and appraising systematic reviews. *Ann Intern Med* 1997;126:532-8.
- 8 Greenhalgh T. How to read a paper. The Medline database. *BMJ* 1997;315:180-3.
- 9 Assendelft WJJ, Tulder MW van, Scholten RJPM, Bouter LM. De praktijk van systematische reviews. II. Zoeken en selecteren van literatuur. *Ned Tijdschr Geneesk* 1999;143:656-61.
- 10 Cochrane database of systematic reviews. In: *The Cochrane Library*. Oxford: Update Software; 1999.
- 11 Assendelft WJJ, Zaat JOM, Meijman FJ. Een Cochrane rubriek in Huisarts en Wetenschap – relevante systematische reviews voor de huisarts. *Huisarts Wet* 2000;43:29-31.
- 12 Jadad AR, Cook DJ, Jones A, Klassen TP, Tugwell P, Moher M, et al. Methodology and reports of systematic reviews and meta-analyses: a comparison of Cochrane reviews with articles published in paper-based journals. *JAMA* 1998;280:278-80.
- 13 Clarke M, Oxman A. Cochrane reviews will be in Medline [letter]. *BMJ* 1999;319:1435.
- 14 Barton S, editor. *Clinical evidence*. 4th ed. Londen: BMJ; 2000.
- 15 Assendelft WJJ, Scholten RJPM, Eijk JThM van, Bouter LM. De praktijk van systematische reviews. III. Methodologische beoordeling van onderzoeken. *Ned Tijdschr Geneesk* 1999;143:714-9.
- 16 Scholten RJPM, Kostense PJ, Assendelft WJJ, Bouter LM. De praktijk van systematische reviews. IV. Het combineren van de resultaten van afzonderlijke onderzoeken. *Ned Tijdschr Geneesk* 1999;143:786-91.
- 17 Scholten RJPM, Assendelft WJJ, Kostense PJ, Bouter LM. De praktijk van systematische reviews. V. Heterogeniteit tussen onderzoeken en subgroepenanalyses. *Ned Tijdschr Geneesk* 1999;143:843-8.
- 18 Assendelft WJJ, Hay EM, Adshear R, Bouter LM. Corticosteroid injections for lateral epicondylitis: a systematic overview. *Br J Gen Pract* 1996;46:209-16.
- 19 Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. *Int J Technol Assess Health Care* 1996;12:195-208.
- 20 Feigin VL, Rinkel GJE, Algra A, Vermeulen M, Gijn J van. Calcium antagonists for aneurysmal subarachnoid haemorrhage (Cochrane Review). In: *The Cochrane Library*. Issue 4. Oxford: Update Software; 1999.
- 21 Mulrow CD, Oxman AD. *Cochrane Collaboration handbook* [updated September 1997]. In: *The Cochrane Library*. Oxford: Update Software; 1999.
- 22 Shea B, Dubé C, Moher D. Assessing the quality of reports of systematic reviews: the QUOROM statement compared to other tools. In: Egger M, Davey Smith G, Altman DG, editors. *Systematic reviews in health care. Meta-analysis in context*. London: BMJ Books; 2001. p. 122-39.
- 23 Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *CMAJ* 1988;138:697-703.
- 24 Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;44:1271-8.
- 25 Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, et al. Agreement among reviewers of review articles. *J Clin Epidemiol* 1991;44:91-8.
- 26 Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of reporting of meta-analyses. *Lancet* 1999;354:1896-900.
- 27 Assendelft WJJ, Koes BW, Knipschild PG, Bouter LM. The relationship between methodological quality and conclusions in reviews of spinal manipulation. *JAMA* 1995;274:1942-8.
- 28 Jadad AR, Cook DJ, Browman GP. A guide to interpreting discordant systematic reviews. *CMAJ* 1997;156:1411-6.
- 29 Cook DJ, Greengold NL, Ellrodt AG, Weingarten SR. The relation between systematic reviews and practice guidelines. *Ann Intern Med* 1997;127:210-6.

Aanvaard op 21 juli 2000