

**FORMULIER Vc**  
**voor het beoordelen van een**  
**SYSTEMATISCHE REVIEW**  
**VAN OBSERVATIONEEL ONDERZOEK**

**Versie maart 2003, geldig tot december 2003**



**Evidence-Based RichtlijnOntwikkeling**



**Formulier Vc: beoordeling systematische review van observationeel onderzoek**

Dit formulier is bestemd voor het beoordelen van systematische reviews van observationeel onderzoek. Dit betreft cohortonderzoek of patiënt-controleonderzoek betreffende bijvoorbeeld etiologische of prognostische vraagstellingen of vraagstellingen betreffende schadelijke bijwerkingen (“harm”).

NB: gelet op de complexiteit van deze materie is het gewenst dat iemand met expertise op het gebied van systematische reviews van observationeel onderzoek bij de beoordeling betrokken is.

Dit formulier is ontwikkeld door een werkgroep bestaande uit vertegenwoordigers van het Dutch Cochrane Centre, het Kwaliteitsinstituut voor de Gezondheidszorg CBO, het Nederlands Huisartsen Genootschap, het institute for Medical Technology Assessment, de Werkgroep Onderzoek Kwaliteit, het College voor Zorgverzekeringen, Zorgonderzoek Nederland (ZonMw) en de Orde van Medisch Specialisten en wordt ondersteund door het Nederlands Paramedisch Instituut, de Vereniging voor Integrale Kankercentra en de Werkgroep Infectieziektenpreventie.

Voor het beoordelen van de kwaliteit van andere typen onderzoek zijn eveneens formulieren ontwikkeld. Deze staan samengevat in onderstaande tabel.

Type onderzoek	Formulier
Dwarsdoorsnedeonderzoek (waarde diagnostische test)	I
Randomised controlled trial (RCT)	II
Cohortonderzoek	III
Patiënt-controleonderzoek	IV
Systematische review van RCT's (therapie en preventie)	Va
diagnostisch onderzoek	Vb
observationeel onderzoek (etiologie/“harm”/prognose)	Vc
Economische evaluatie	VI
Richtlijn	AGREE

**Instructie beoordeling**

- De bruikbaarheid van een publicatie voor een richtlijn wordt in de formulieren op drie facetten beoordeeld: validiteit, toepasbaarheid in de praktijk en toepasbaarheid in de Nederlandse gezondheidszorg
- Daarnaast wordt gevraagd om de belangrijkste kwantitatieve gegevens te extraheren en op een uniforme wijze te presenteren.
- De opmaak van de beoordelingsformulieren maakt het u makkelijk:
  - a) op diverse plaatsen is een beslismoment ingebouwd: indien een publicatie op dat moment niet aan de vereisten van validiteit of toepasbaarheid voldoet hoeft u met de beoordeling niet verder te gaan.
  - b) de criteria en manier van data-extractie worden telkens op de tegenoverliggende pagina kort toegelicht.

Zend opmerkingen of suggesties aangaande dit formulier naar [cochrane@amc.uva.nl](mailto:cochrane@amc.uva.nl).

Vraag 1. *Vraagstelling.* Een systematische review dient uit te gaan van een expliciete vraagstelling. De vraag dient bij voorkeur de volgende aspecten te bevatten: de doelpopulatie, de onderzochte etiologische of prognostische factor of interventie en de onderzochte uitkomst (ziekte of schadelijke bijwerking). Een leidraad kan het Engelse acroniem "PICO" zijn: **P**atient, **I**ntervention ("interventie" hier op te vatten als de onderzochte factor), **C**omparison (wordt in dit kader meestal niet benoemd) en **O**utcome.

Vraag 2. *Zoekactie.* Uitgaande van de vraagstelling dient op gestructureerde wijze met behulp van duidelijk omschreven zoektermen in de internationale literatuur te worden gezocht. De zoektermen moeten zodanig gekozen zijn, dat naar verwachting nagenoeg alle relevante artikelen op het omschreven gebied gevonden zullen worden. Een minimum vereiste omtrent de databases is een zoekactie in MEDLINE en controle van de referenties van de verzamelde artikelen. Indien niet in MEDLINE gezocht is, hoeft u met de beoordeling niet verder te gaan. Of een aanvullende zoekactie in EMBASE of specialistische databases zoals Cinahl en Psyclit meerwaarde heeft bij het zoeken naar observationeel onderzoek, is niet bekend. Optionele aanvullende zoekacties zijn gedrukte indexen, het aanschrijven van experts en fabrikanten en het handmatig doorzoeken van tijdschriften. Eveneens zullen de eventuele restricties in de zoekactie beschreven moeten zijn. Hierbij kan gedacht worden aan restrictie in jaartal van publicatie, publicatiestatus (wel of niet gepubliceerd) of in taal. Hoewel geen empirisch onderzoek gedaan is naar de invloed van taalrestricties op het resultaat van een review van observationeel onderzoek (bijvoorbeeld alleen inclusie van Engelstalige onderzoeken) kan een beperking tot één of meerdere talen leiden tot minder precieze, en mogelijk ook vertekende, schattingen.

Vanzelfsprekend moet de zoekactie aansluiten op de vraagstelling zoals beoordeeld in vraag 1.

Vraag 3. *Selectie.* Een uitputtende zoekactie (vraag 2) zal resulteren in een groot aantal treffers. Veel aldus gevonden artikelen zullen niet relevant zijn voor de systematische review en dus buiten de systematische review gelaten moeten worden. De selectie van artikelen die geïncludeerd moeten worden, dient bij voorkeur door twee reviewers onafhankelijk van elkaar te geschieden aan de hand van expliciet beschreven in- en exclusiecriteria. Een minimum van twee reviewers voor de definitieve selectie is bedoeld om de kans op selectiebias in te perken.

Onder in- en exclusiecriteria vallen bijvoorbeeld: een specifieke onderzoekspopulatie (bijvoorbeeld personen in de open bevolking of patiënten in die naar de fysiotherapeut zijn verwezen), gebruik van een bepaalde interventie (bijvoorbeeld manipulatie), bestudering van een bepaalde determinant (bijvoorbeeld alcoholgebruik), etc. De in- en exclusiecriteria dienen aan te sluiten op de vraagstelling en dienen expliciet beschreven te zijn om de reproduceerbaarheid van de systematische review te verhogen. Tevens moet de selectie vanaf de initiële database zoekactie tot de uiteindelijke inclusie in de systematische review duidelijk beschreven zijn.

Vraag 4. *Kwaliteitsbeoordeling.* De kwaliteit van de opgenomen onderzoeken dient beoordeeld te zijn. Indien geen kwaliteitsbeoordeling heeft plaatsgevonden, hoeft u niet met de beoordeling verder te gaan. Kwaliteitsbeoordeling is bij voorkeur uitgevoerd door twee reviewers onafhankelijk van elkaar aan de hand van een expliciete lijst van criteria. De onafhankelijke beoordeling verkleint de kans op informatiebias.

De kwaliteit van de opgenomen onderzoeken is mede bepalend voor de validiteit van de conclusies van en aanbevelingen in de systematische review. Er zijn verschillende criterialijsten om de kwaliteit van observationeel onderzoek te beoordelen, maar gevalideerde checklists zijn niet voorhanden.

Onderzoek heeft laten zien dat een associatie kan worden overschat wanneer onvoldoende is gecorrigeerd voor versturende factoren en wanneer (bij etiologisch onderzoek en onderzoek naar bijwerkingen) sprake is van een patiënt-controleonderzoek. Patiënt-controleonderzoek is gevoeliger voor selectiebias en voor informatiebias dan prospectief cohortonderzoek. Bij systematische reviews van prognostisch onderzoek is het belangrijk dat de volledigheid van follow-up is meegenomen in de beoordeling van geselecteerde onderzoeken.

De reviewers dienen duidelijk beschreven te hebben op welke aspecten zij de oorspronkelijke onderzoeken beoordeeld hebben en het resultaat van deze kwaliteitsbeoordeling dient per individueel onderzoek in de review gerapporteerd te zijn.

De wijze waarop tot consensus gekomen wordt en de mate van overeenstemming tussen de reviewers dient beschreven zijn.

## Beoordeling van de kwaliteit van een systematische review van observationeel onderzoek

Naam beoordelaar: ..... Datum: .....

Titel: .....

Auteurs: .....

Bron: .....

## Beoordeling van de validiteit

Korte beschrijving van de blootstelling of prognostische factor(en): .....

.....

## METHODEN

1. Is de vraagstelling adequaat geformuleerd?  
 Ja  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden
  
2. Is de zoekactie adequaat uitgevoerd?  
 Ja  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden
  
3. Is de selectieprocedure van artikelen adequaat uitgevoerd?  
 Ja  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden
  
4. Is de kwaliteitsbeoordeling adequaat uitgevoerd?  
 Ja  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden

Vraag 5. *Data-extractie*. Data-extractie is het verzamelen van de resultaten uit de oorspronkelijke onderzoeken. In geval van observationeel onderzoek betreft dit aspecten van de bestudeerde blootstelling of prognostische factor (bijvoorbeeld welke eenheid van blootstelling het onderzoek bestudeerde), informatie over welke potentiële confounders in de oorspronkelijke onderzoeken beschouwd waren en het voor deze confounders gecorrigeerde effect per eenheid van blootstelling. Soms dienen de reviewers gegevens zelf te reconstrueren. Daarom dient het data-extractieproces duidelijk beschreven te zijn en weer bij voorkeur door twee reviewers onafhankelijk van elkaar plaatsgevonden te hebben.

Vraag 6. *Beschrijving oorspronkelijke onderzoeken*. De belangrijkste kenmerken van de oorspronkelijke onderzoeken moeten in de systematische review beschreven zijn. Wat belangrijk is, hangt af van het onderwerp. Tenminste beschreven moeten zijn:

- *onderzochte populatie*: per onderzoek tenminste de manier van selectie (aard en oorsprong van de onderzoekspopulatie bijvoorbeeld “aselecte steekproef uit de open bevolking aan de hand van het bevolkingsregister”), tijdsperiode, leeftijds- en geslachtsverdeling; in geval van prognostisch onderzoek en onderzoek naar schadelijke bijwerkingen: gegevens over de ziekte zoals aard, ernst en stadium
- *blootstelling*: definitie, eenheid, wijze van vaststelling; in geval van onderzoek naar schadelijke bijwerkingen: gegevens over de onderzochte interventie (dosis, duur, wijze van toediening, etc.)
- *uitkomsten (ziekte / schadelijke bijwerking)*: definitie, wijze van vaststelling
- *confounders*: aard en eenheid
- *resultaten*: voor alle relevante confounders gecorrigeerde effect per eenheid van de determinant (blootstelling of prognostische factor)

Vraag 7. *Meta-analyse*. Indien de geselecteerde onderzoeken voldoende vergelijkbaar zijn, kunnen de resultaten worden gecombineerd tot één algemene schatting van de sterkte van de bestudeerde associatie. Bij observationeel onderzoek zal echter vaak sprake zijn van grote verschillen in onderzoeksresultaten (heterogeniteit). Deze verschillen kunnen door toeval verklaard worden of door werkelijke verschillen tussen onderzoeken, zoals verschillen in onderzoeksontwerp (cohort- of patiënt-controleonderzoek), onderzoekspopulatie, wijze van vaststelling van de determinant of definitie van de uitkomst. Een systematische review van observationeel onderzoek dient de invloed van deze mogelijke bronnen van methodologische of klinische heterogeniteit op de uitkomst dan ook te onderzoeken.

Onderzoek van bronnen van heterogeniteit geschiedt aan de hand van subgroepanalyses of meta-regressieanalyse met de afzonderlijke onderzoeken als analyse-eenheid. Doorgaans gaat een voorbereiding van de resultaten van de afzonderlijke onderzoeken vooraf aan deze analyses. Deze voorbereiding betreft met name de eenheid van blootstelling en dient om de oorspronkelijke onderzoeken althans op dit punt beter vergelijkbaar te maken.

Let bij de beoordeling van een meta-analyse van observationeel onderzoek op de volgende aspecten.

1. Beoordeel of de auteurs alle relevante potentiële bronnen van klinische heterogeniteit (= potentiële subgroepen) hebben beschouwd. De in aanmerking komende bronnen van heterogeniteit dienen bij voorkeur vooraf benoemd te zijn. Bronnen van heterogeniteit die aan de hand van zogenoemde *post-hoc* analyses gedetecteerd zijn, dienen met grote voorzichtigheid beoordeeld te worden, aangezien er een grote kans is op het vinden van fout-positieve resultaten.
2. De auteurs dienen op de een of andere wijze de mate van heterogeniteit tussen de onderzoeksresultaten beschouwd te hebben. Dit kan door visuele inspectie van de afzonderlijke effectschattingen met hun betrouwbaarheidsinterval (is er voldoende overlap?) of door een formele statistische toets (in dit kader vaak Q-toets genoemd).
3. Subgroepanalyses en meta-regressieanalyses dienen bij voorkeur aan de hand van het zogenoemde “random effects model” uitgevoerd te zijn. In dit model wordt rekening gehouden met de (rest)variatie tussen onderzoeken (= heterogeniteit).

In voorkomende gevallen (bijvoorbeeld als de oorspronkelijke onderzoeken van zeer slechte kwaliteit zijn of te veel uiteenlopen wat betreft onderzoekskenmerken) kan ook terecht afgezien zijn van pooling. In dat geval kan worden volstaan met een kwalitatieve samenvatting van de resultaten. Dit geldt ook voor systematische reviews van prognostisch onderzoek dat gericht is op de vraag welke combinatie van factoren het beste het beloop van een ziekte kan voorspellen (prognostische modellen). Poolen van de resultaten van dergelijke modellen is doorgaans niet goed mogelijk.

5. Is adequaat beschreven hoe data-extractie heeft plaatsgevonden?  
 Ja  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden
6. Zijn de belangrijkste kenmerken van de oorspronkelijke onderzoeken beschreven?  
 Ja  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden
7. Is meta-analyse op een correcte manier uitgevoerd?  
 Ja / niet van toepassing (terecht geen meta-analyse uitgevoerd)  
 Nee  
 Te weinig informatie in het artikel om dit te beantwoorden

#### ALGEMEEN OORDEEL

8. Zijn de resultaten van de systematische review valide en toepasbaar?  
 Voldoende valide en toepasbaar → ga verder bij 9  
 Twijfelachtig → ga verder bij 9; voorzichtig oordeel bij 12  
 Onvoldoende valide en toepasbaar U kunt stoppen met het invullen van de checklist, tenzij er geen betere artikelen op dit gebied zijn (terugkoppelen naar de werkgroep)

Vraag 8. *Algemeen oordeel.* Hier wordt een inschatting gevraagd van de validiteit (Is dit een goede systematische review?) en toepasbaarheid (Gaat deze review over mijn PICO?). Let hierbij ook op eventuele fouten in de review die funest zijn voor de validiteit ervan (*red flags, fatal flaws*). Er zijn geen regels te geven voor welke items positief gescoord moeten worden of hoeveel items tenminste positief gescoord moeten worden. Dit is deels afhankelijk van de “state-of-the-art” met betrekking tot het betreffende onderwerp. Het gaat er hier om het samenvattend oordeel van wat de beoordelaar de werkgroep zou willen mededelen over de bruikbaarheid van het artikel voor de besluitvorming.

Vraag 9. *Resultaten.*

NB: zie ook de toelichting bij vraag 7!

Aan de hand van subgroepanalyses en meta-regressieanalyses worden schattingen verkregen van het effect van blootstelling voor verschillende subgroepen. Neem, indien dergelijke pooling is uitgevoerd, de door de auteurs gepresenteerde subgroep-specifieke gepoolde resultaten over. (NB: U bent hier afhankelijk van de reviewers en kunt niet zelf uw eigen favoriete effectmaat berekenen, tenzij alle gegevens gerapporteerd zijn en u de hele analyse zelf kunt herhalen.)

Voor dichotome uitkomsten (ziek / niet-ziek; wel / geen bijwerking) noteert u de gepoolde relatieve risico's (RR), de gepoolde oddsratio's (OR) of de gepoolde absolute risico reducties (ARR), voor continue variabelen (numerieke waarden) de gepoolde verschillen (Engels: mean difference (MD)) of de gepoolde gestandaardiseerde verschillen (Engels: standardised mean difference (SMD)).

Bij deze associatiematen (effectmaten) wordt een 95%-betrouwbaarheidsinterval (95%-BI) gerapporteerd. Een (gepoold) resultaat is significant indien het 95%-BI de “neutrale” waarde van de desbetreffende effectmaat niet omvat. De neutrale waarde is de waarde waarvoor geen verschil in effect bestaat tussen de onderzochte interventies. In geval van OR of RR is de neutrale waarde “1” (het geschatte effect is dus “significant” als het 95%-BI de “1” niet omvat). In geval van ARR of een (gestandaardiseerde) verschilscore is de neutrale waarde “0” (het geschatte effect is dan “significant” als het 95%-BI de “0” niet omvat).

In het geval van een systematische review van cohortonderzoek naar bijwerkingen is het zinvol een number needed to harm (NNH) te berekenen voor verschillende achtergrond risico's (kans op de uitkomst in de niet-blootgestelde groep). In de appendix is een tabel opgenomen voor het afleiden van het NNH van oddsratio's en relatieve risico's voor verschillende achtergrond risico's. (NB: het achtergrond risico dient uitgedrukt te worden als proportie en niet als percentage!)

9. Resultaten:

Uitkomst: .....

Eenheid van blootstelling: .....

Confounders in de analyse: .....

Follow-up: ..... weken / maanden / jaar

	Waarde (95%-BI)			
Gecorrigeerd gepoold effect per eenheid van blootstelling	Subgroep ..	Subgroep ..	Subgroep ..	Subgroep ..
Oddsratio (OR)				
Relatieve risico (RR)				
Absolute risicoreductie (ARR)				
(Gestandaardiseerd) verschil van gemiddelden ((S)MD)				
Variabele follow-up duur:				
Hazard ratio (HR)				

Number needed to harm (zie toelichting)

Achtergrondrisico	Number needed to harm (NNH) (zie tabel en/of formules in de appendix)			

Vraag 10 en 11. *Toepasbaarheid in de Nederlandse gezondheidszorg*. Beide vragen zijn een belangrijk onderdeel van richtlijnontwikkeling en dienen daarom in de werkgroep bediscussieerd te worden.

Vraag 12. *Conclusie met betrekking tot de systematische review en het effect van de blootstelling of de prognostische factor*

Geef hier een globale samenvatting van het eindoordeel over de systematische review. Probeer, indien aanwijzingen bestaan voor vertekening van de resultaten, tenminste een inschatting te maken van de richting van de vertekening (overschatting of onderschatting van het effect) en zo mogelijk ook over de grootte van de vertekening. Eventuele aanwijzingen voor mogelijke belangenverstremming van de auteurs met belanghebbende opdrachtgevers, kunt u hier ook rapporteren. Ook is het verstandig ingezonden brieven en/of redactionele commentaren op het hier door u beoordeelde onderzoek te raadplegen bij het formuleren van uw conclusie.

*Voorbeeld:* "Uitstekend uitgevoerde review door onafhankelijke reviewers. Uitgangspunt is een duidelijk vraagstelling met een duidelijke definitie van blootstelling. Er is een uitgebreide zoektocht uitgevoerd zonder restricties aan de hand van adequaat gekozen zoektermen. De opgenomen onderzoeken zijn voor het merendeel van redelijke kwaliteit. Blootstelling aan factor X verhoogt de kans op uitkomst Y bij mannen maar niet bij vrouwen. Roken lijkt geen duidelijke effect-modificator te zijn van de relatie tussen X en Y. Het is zinvol om manlijke patiënten op de vergrote risico's van X te wijzen."

## TOEPASBAARHEID IN DE NEDERLANDSE GEZONDHEIDSZORG

10. Kan het gevonden resultaat worden toegepast op de Nederlandse situatie?  
(hierbij valt bijvoorbeeld te denken aan het voorkomen van de onderzochte factor)

- Ja
- Nee
- Te weinig informatie in het artikel om dit te beantwoorden

11. Op welk(e) echelon(s) kan het resultaat worden toegepast?  
(meerdere opties tegelijk mogelijk)

- algemene bevolking
- eerste lijn
- tweede lijn
  - academische ziekenhuizen
  - perifere ziekenhuizen
- derde lijn

## CONCLUSIE

12. Conclusie met betrekking tot de systematische review

**APPENDIX**

Tabel A. "Numbers needed to harm" naar achtergrondrisico (= kans op de uitkomst in de niet-blootgestelde groep = "control event rate" = CER) en (gecorrigeerde) relatieve risico (RR).

Formule:  $NNT = 1 / [CER * (RR - 1)]$

Achtergrondrisico	Relatieve risico (RR)								
	1,1	1,25	1,5	1,75	2,0	2,25	2,5	3,0	4,0
0,001	10000	4000	2000	1334	1000	800	667	500	334
0,005	2000	800	400	267	200	160	134	100	67
0,010	1000	400	200	134	100	80	67	50	34
0,015	667	267	134	89	67	54	45	34	23
0,02	500	200	100	67	50	40	34	25	17
0,05	200	80	40	27	20	16	14	10	7
0,10	100	40	20	14	10	8	7	5	4
0,15	67	27	14	9	7	6	5	4	3
0,20	50	20	10	7	5	4	4	3	2

Tabel B. "Numbers needed to harm" naar achtergrondrisico (= kans op de uitkomst in de niet-blootgestelde groep = "control event rate" = CER) en (gecorrigeerde) oddsratio (OR).

Formule:  $NNT = 1 / [CER * OR / (1 - CER) / (1 + CER * OR / (1 - CER)) - CER]$

Achtergrondrisico	Oddsratio (OR)								
	1,1	1,25	1,5	1,75	2,0	2,25	2,5	3,0	4,0
0,001	10012	4006	2004	1336	1003	802	669	502	335
0,005	2012	806	404	270	203	162	136	102	69
0,010	1012	406	204	136	103	82	69	52	35
0,015	678	272	137	92	69	56	47	35	24
0,02	512	206	104	70	53	42	36	27	19
0,05	212	86	44	30	23	18	16	12	9
0,10	113	46	24	16	13	10	9	7	5
0,15	80	33	17	12	10	8	7	6	4
0,20	64	27	14	10	8	7	6	5	4