

FORMULIER III
voor het beoordelen van een
COHORTONDERZOEK

Versie februari 2003, geldig t/m december 2003



Evidence-Based RichtlijnOntwikkeling

Formulier III: beoordeling cohortonderzoek

Dit formulier is bestemd voor het beoordelen van cohortonderzoek. Cohortonderzoek wordt uitgevoerd om etiologische of prognostische vraagstellingen te onderzoeken of bijwerkingen van een bepaalde interventie.

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 Specialist 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
observatieonderzoek (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 en suggesties aangaande dit formulier naar cochrane@amc.uva.nl.

In cohortonderzoek wordt een afgebakende groep patiënten gedurende een bepaalde tijd gevolgd en wordt gekeken of een bepaalde uitkomst optreedt. Cohortonderzoek kan niet alleen vraagstellingen betreffen op etiologisch gebied of op het gebied van bijwerkingen ("Harm"), maar ook prognostische vraagstellingen. In cohortonderzoek wordt de relatie tussen een determinant (een etiologische factor, een bepaalde interventie of een prognostische factor) en de uitkomst onderzocht. Hierbij wordt een vergelijking gemaakt tussen een groep deelnemers met een bepaalde determinant en een groep deelnemers zonder deze determinant. In deze beoordelingslijst zal gesproken worden over blootgestelde deelnemers en niet-blootgestelde deelnemers aan een bepaalde determinant (een etiologische of een prognostische factor of een interventie).

Cohortonderzoeken kunnen prospectief zijn (de blootstelling is gedefinieerd en de deelnemers zijn geselecteerd voordat de uitkomst optrad) of retrospectief (historisch cohortonderzoek; de blootstelling is bekeken, meestal in medische statussen, nadat de uitkomst al bekend was). Retrospectieve cohortonderzoeken zijn gevoeliger voor vertekening.

In cohortonderzoek kan sprake zijn van twee te onderscheiden situaties:

- 1) De relatie tussen één centrale determinant en de uitkomst wordt onderzocht. Deze situatie komt voor in etiologisch onderzoek of onderzoek naar bijwerkingen. Soms is men bij prognostisch onderzoek geïnteresseerd in slechts één determinant. Kenmerk van deze situaties is, dat een schatting van de sterkte van de relatie tussen de determinant en de uitkomst wordt gemaakt die gecorrigeerd wordt voor allerlei mogelijke andere factoren (confounders).
- 2) De relatie tussen meerdere determinanten en de uitkomst wordt onderzocht. Deze methode wordt met name toegepast in prognostisch onderzoek. Hierbij wordt een prognostisch model geschat en wordt gekeken wat de invloed is van die verschillende determinanten op de uitkomst.

Dit onderscheid is met name van belang bij het beoordelen van de vragen 8 en 10.

Vraag 1. Definitie van de studiegroepen. De belangrijkste kenmerken van de blootgestelde en niet-blootgestelde groep dienen duidelijk beschreven te zijn. Bij prognostisch onderzoek dient er sprake te zijn van een cohort dat is samengesteld uit patiënten op eenzelfde tijdstip in het beloop van hun ziekte (inceptiecohort).

Vraag 2. Selectiebias. Om te beoordelen of sprake is van selectiebias moeten de inclusie- en exclusiecriteria van de onderzoeksgroepen duidelijk beschreven zijn. Selectiebias kan bijvoorbeeld ontstaan als retrospectief een cohort wordt samengesteld na kennis over de uitkomst. De beslissing om deelnemers te includeren kan dan afhankelijk zijn van de uitkomst. Hierdoor wordt de kans om als "zieke" blootgestelde geselecteerd te worden in het cohort verschillend van de kans om als "zieke" niet-blootgestelde geselecteerd te worden. Selectiebias kan ook ontstaan wanneer er maar een deel van het oorspronkelijk cohort geselecteerd wordt, bijvoorbeeld op basis van een onbewuste voorkennis over een relatie tussen bepaalde factoren en de uitkomst. Hierdoor kunnen belangrijke verschillen ontstaan tussen de te vergelijken onderzoeksgroepen (naast de blootstelling) die samenhangen met de uiteindelijke uitkomst waarvoor niet gecorrigeerd kan worden.

Vraag 3. Blootstelling. De blootstelling dient duidelijk gedefinieerd te zijn en er moet aangeven zijn met welk instrument, wanneer en onder welke omstandigheden de determinant is vastgesteld. Dit dient in beide onderzoeksgroepen op dezelfde wijze gedaan te zijn.

Vraag 4. Uitkomst. De uitkomst en gebruikte criteria voor het vaststellen van de uitkomst dienen duidelijk gedefinieerd te zijn, zodanig dat dit door andere onderzoekers herhaald kan worden. De uitkomst dient op een valide manier gemeten te zijn.

Vraag 5. Geblindeerde uitkomstmeting. Blinding wordt bereikt door zonder kennis over de blootstellingstatus de uitkomst te meten. Door de blinding wordt voorkomen dat de uitkomstmeting beïnvloed wordt door kennis over de blootstelling. Indien er geen blinding heeft plaatsgevonden zijn er dan aanwijzingen over de invloed van deze kennis over de blootstellingstatus op de beoordeling van de uitkomst?

Beoordeling van de kwaliteit van een cohortonderzoek

Naam beoordelaar: Datum:

Titel:

Auteurs:

Bron:

Beoordeling van de validiteit

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

.....

.....

VALIDITEIT

1. Zijn de te vergelijken onderzoeksgroepen duidelijk gedefinieerd?

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

2. Kan selectiebias voldoende worden uitgesloten?

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

3. Is de blootstelling duidelijk gedefinieerd en is de methode voor beoordeling van blootstelling adequaat?

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

4. Is de uitkomst duidelijk gedefinieerd en is de methode voor beoordeling van de uitkomst adequaat?

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

5. Is de uitkomst blind voor de blootstellingstatus bepaald?

 Ja Nee → is dit van invloed op beoordeling van de uitkomst? Ja Nee Te weinig informatie in het artikel om dit te beantwoorden

Vraag 6. *Follow-up*. De follow-up duur dient zodanig lang te zijn dat de bestudeerde uitkomst erin op kan treden.

Vraag 7. *Selectieve loss-to-follow-up*. Er dienen geen belangrijke verschillen in uitval van deelnemers tussen de onderzoeksgroepen te zijn. Aantallen en redenen voor uitval dienen gerapporteerd te zijn. Ook als er geen uitvallers waren, moet dit zijn beschreven. Het is niet mogelijk om op voorhand per indicatiegebied aan te geven welk percentage uitval nog acceptabel is. Relatief grote uitval maakt een cohortonderzoek gevoelig voor selectieve uitval. Dit kan het geval zijn als de absolute aantallen uitvallers en de redenen van uitval tussen de onderzoeksgroepen verschillen.

Vraag 8. *Confounders / prognostische factoren*.

a) *Onderzoek naar het effect van één centrale determinant*:

Confounding is een verstoring van een relatie tussen blootstelling en uitkomst door een andere factor die zowel met de blootstelling als de uitkomst geassocieerd is. In cohortonderzoek zal vrijwel altijd sprake zijn van een ongelijke verdeling van confounders over de groepen. Klinisch inzicht zal moeten helpen om te beoordelen of alle belangrijke confounders in overweging zijn genomen. Er dient te zijn aangegeven welke confounders er overwogen zijn en hoe daar in het ontwerp van het onderzoek en in de analyse rekening mee is gehouden. In de ontwerpfasen kan dit door restrictie of matching, in de analysefase door gestratificeerde analyse of multivariate analyse. Wanneer geslacht bijvoorbeeld een mogelijke confounder is, kan ervoor gekozen worden om het onderzoek louter te richten op mannen (restrictie) of evenveel mannen te laten deelnemen in de blootgestelde als in de niet-blootgestelde groep (matching). Ook kan achteraf in de analyse gecorrigeerd worden voor "geslacht" door een gestratificeerde analyse uit te voeren of een multivariate analyse waarin geslacht meegenomen wordt (vraag 10).

Een voorbeeld van confounding: er werd een relatie gevonden tussen pilgebruik en het optreden van een myocardinfarct. Later echter bleek dat deze relatie mogelijk te wijten was aan een hoger percentage rokers in de pilgroep. Toen voor "roken" in de analyse gecorrigeerd werd, bleek de relatie tussen pilgebruik en myocardinfarct niet meer waarneembaar. Roken, een determinant van myocardinfarct, is in dit voorbeeld een confounder van de relatie tussen pilgebruik en myocardinfarct.

b) *Onderzoek betreffende meerdere determinanten tegelijk (prognostische modellen)*:

In het geval van prognostisch onderzoek waarbij onderzocht wordt of meerdere factoren een uitkomst voorspellen is het van belang dat alle relevante prognostische factoren meegenomen zijn in de analyse.

Vraag 9. *Algemeen oordeel*. Hier wordt een inschatting gevraagd van de validiteit (Is dit een goed onderzoek?) en toepasbaarheid (Gaat dit onderzoek over mijn PICO?). Let hierbij ook op eventuele fouten in het onderzoek 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.

6. Is er een voldoende lange follow-up?
 Ja
 Nee
 Te weinig informatie in het artikel om dit te beantwoorden
7. Kan selectieve loss-to-follow-up voldoende worden uitgesloten?
 Ja
 Nee
 Te weinig informatie in het artikel om dit te beantwoorden / loss-to-follow-up niet beschreven
8. Zijn de belangrijkste confounders of prognostische factoren geïdentificeerd en is er adequaat rekening mee gehouden in het ontwerp van het onderzoek of in de analyse?
 Ja
 Nee
 Te weinig informatie in het artikel om dit te beantwoorden

TUSSENOORDEEL

9. Zijn de resultaten van het onderzoek valide en toepasbaar?
 Voldoende valide en toepasbaar → ga verder bij 10
 Twijfelachtig → ga verder bij 10
 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 10. Resultaten

a) Onderzoek naar het effect van één centrale determinant:

Noteer de bestudeerde uitkomst en de eenheid van blootstelling waarvoor het effect berekend is (bijvoorbeeld gebruik van antihistaminica versus geen gebruik, per 10 sigaretten per dag). Noteer ook de confounders die in de multivariate analyse (zie hierna) zijn opgenomen, en de follow-up duur.

De uitkomsten van cohortonderzoek kunnen dichotoom zijn (wel/geen schade; wel/geen bijwerking; wel/geen genezing; wel/geen overleving) of continu (bloeddruk; cholesterolwaarde). Bij dichotome uitkomsten kan het effect van één eenheid blootstelling uitgedrukt worden in een relatief risico (RR), oddsratio (OR) of absolute risico reductie (ARR). Bij continue uitkomsten wordt het effect uitgedrukt in een verschilscore (Engels: mean difference (MD)). Indien sprake is van een variabele follow-up duur wordt survival-analyse uitgevoerd. Naast de hazardratio (zie hierna) kan tevens de mediane duur tot het optreden van de uitkomst bepaald worden (bijvoorbeeld mediane duur tot recidief).

In alle bovengenoemde gevallen dient een analyse uitgevoerd te zijn waarmee "gecorrigeerd" is voor potentiële confounders (prognostische factoren die ongelijk verdeeld zijn over beide onderzoeksgroepen). Dit zal bij dichotome uitkomsten doorgaans multivariate logistische regressie of survival-analyse (Cox regressie) betreffen en bij continue uitkomsten multivariate lineaire regressie. De uitkomst van dergelijke regressieanalyses is een schatting van het effect, die gecorrigeerd is voor alle confounders die in het regressiemodel zijn opgenomen. Logistische regressie levert een gecorrigeerde OR, survival-analyse een gecorrigeerd RR (in deze context hazard ratio (HR) genoemd) en lineaire regressie een gecorrigeerde verschilscore.

Een ander type analyse waarmee gecorrigeerd kan worden voor confounding, is gestratificeerde analyse, bijvoorbeeld de methode van Mantel-Haenszel. Hierin wordt echter steeds slechts voor 1 confounder gecorrigeerd. Hiermee kunnen gecorrigeerde OR's, RR's of ARR's verkregen worden.

Noteer het voor confounding gecorrigeerde effect ("adjusted" OR, RR, ARR of MD) met 95%-betrouwbaarheidsinterval. (NB: U bent hier afhankelijk van de auteurs en kunt doorgaans zelf geen resultaten berekenen.)

Probeer in het geval van onderzoek betreffende schadelijke bijwerkingen ("Harm") nog het "number needed to harm" (NNH) te berekenen voor verschillende achtergrondrisico's (= kans op de bijwerking in de niet-blootgestelde groep). In de appendix is een tabel opgenomen voor het afleiden van NNH's van oddsratio's en relatieve risico's voor verschillende achtergrondrisico's. (NB: het achtergrondrisico dient uitgedrukt te worden als proportie en niet als percentage!)

b) Prognostisch onderzoek:

In prognostisch onderzoek worden de in het onderzochte cohort geobserveerde uitkomsten ("observed frequency") ook wel vergeleken met de resultaten zoals die in een fictief extern cohort zouden zijn opgetreden ("expected frequency"). Deze expected frequency kan bijvoorbeeld verkregen worden door aan de hand van CBS-gegevens betreffende de algemene bevolking te berekenen hoeveel uitkomsten opgetreden zouden zijn als het externe cohort precies dezelfde leeftijds- en geslachtsopbouw zou hebben gehad als het bestudeerde cohort. De observed frequency wordt dan afgezet tegen de expected frequency waarmee een "standardized morbidity ratio" (of in geval van de uitkomst "overleden" een "standardized mortality ratio") verkregen wordt. Deze maat is als een relatief risico op te vatten, waarbij alleen gecorrigeerd is voor leeftijd en geslacht.

Indien een prognostisch model is geschat, noteer dan welke prognostische factoren de uitkomst beïnvloeden en hoe sterk zij dit doen (dit laatste zal meestal uitgedrukt zijn als OR's of als HR's).

10. Resultaten

a) *Eén centrale determinant:*

Uitkomst:

Eenheid van blootstelling:

Confounders in de analyse:

Follow-up: weken / maanden / jaar

Gecorrigeerd effect per eenheid van blootstelling	Waarde	95%-betrouwbaarheidsinterval
Oddsratio (OR)		
Relatieve risico (RR)		
Absolute risico reductie (ARR)		
Verskil van gemiddelden (MD)		
<i>Variabele follow-up duur:</i>		
Hazard ratio (HR)		

Achtergrondrisico	Number needed to harm (NNH) (zie tabel in appendix)

b) *Prognostisch onderzoek:*

Uitkomst:

Follow-up: weken / maanden / jaar

	Waarde	95%-betrouwbaarheidsinterval
Standardized mortality or morbidity ratio (SMR)		

Prognostisch model (factoren)	Waarde	95%-betrouwbaarheidsinterval

Vraag 11 en 12. *Toepasbaarheid in de Nederlandse gezondheidszorg*. Beide vragen zijn een belangrijk onderdeel van richtlijnontwikkeling en dienen daarom in de werkgroep bediscussieerd te worden.

Vraag 13. *Conclusie met betrekking tot het artikel*

Geef hier een globale samenvatting van het eindoordeel over het artikel. Probeer, indien aanwijzingen bestaan voor vertekening van de resultaten, tenminste een inschatting te maken van de richting van de vertekening (overschatting of onderschatting van de bestudeerde associatie) 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: "Eindoordeel voldoende. Goed opgezet artikel. De onderzoekspersonen zijn echter maar een jaar gevolgd. Over bijwerkingen op de lange termijn kan geen uitspraak worden gedaan. Middel A lijkt minder hoofdpijnklachten en duizelingen te geven dan middel B bij gelijke effectiviteit. Daarnaast dient een afweging gemaakt te worden tussen bijvoorbeeld de gevonden nadelen van een middel ten opzichte van de voordelen en moeten bijvoorbeeld de mogelijkheden tot preventie in overweging genomen worden."

TOEPASBAARHEID IN DE NEDERLANDSE GEZONDHEIDSZORG

11. Kan het gevonden resultaat vertaald worden op de Nederlandse situatie?

Ja

Nee

Te weinig informatie in het artikel om dit te beantwoorden

12. 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

13. Conclusie met betrekking tot het artikel

Appendix

Tabel A. Cohortonderzoek naar bijwerkingen: “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