

Kritisch lezen

Een bespreking van een onderzoeksverslag aan de hand van het Draaiboek „kritisch lezen” van de Commissie Wetenschappelijk Onderzoek.

JOAN BOEKE

Enkele maanden geleden is in dit tijdschrift uitvoerig aandacht geschonken aan de CWO-publikatie *Wetenschappelijk onderzoek in de huisartsgeneeskunde*.^{1 2} Waarschijnlijk is het *Draaiboek „kritisch lezen”* echter een veel bekender loot aan de stam van deze vruchtbare commissie: velen hebben de afgelopen jaren kennis kunnen maken met de zakelijke en doelgerichte aanpak die in dit draaiboek is neergelegd. Het was dan ook bijzonder teleurstellend voor de leden van de CWO, dat het geruime tijd niet mogelijk bleek een verslag van een dergelijke „kritisch lezen”-sessie gepubliceerd te krijgen. Gelukkig deed zich recent een ongezochte gelegenheid voor, toen drie CWO-leden *Huisarts en Wetenschap* een artikel aanboden over een onderzoek in de eigen praktijk.³ Op verzoek van de redactiecommissie hebben de overige leden van de commissie zich „kritisch” over het eerste deel van dit artikel gebogen en de volgende bijdrage is daarvan het resultaat.

Inleiding

In deze bijdrage zal verslag worden gedaan van een bespreking na het „kritisch lezen” van een onderzoeksartikel. Dat kritisch lezen is gebeurd door de leden van de Commissie Wetenschappelijk Onderzoek aan de hand van een aantal vragen. De CWO heeft deze vragen alsmede een *Draaiboek „kritisch lezen”* opgesteld als onderdeel van een onderwijsprogramma Wetenschappelijk Onderzoek in de Huisartsgeneeskunde. De inhoudelijke materie van dit programma is neergelegd in een boek dat in oktober 1983 door de CWO is gepubliceerd.¹

De bedoeling van de Cursus „kritisch lezen” is om huisartsen (in opleiding) vertrouwd te maken met een aantal belangrijke onderzoeksbegrippen door het bieden van richtlijnen waarmee systematisch een onderzoeksartikel kan worden gelezen en beoordeeld. Er wordt daarbij gerefereerd aan de tekst van het hierboven genoemde boek.

De CWO-leden hebben de afgelopen jaren veel groepen begeleid waarin een artikel „kritisch” werd gelezen en besproken volgens het draaiboek. Plaatselijke NHG-groepen, waarneemgroepen,

workshops tijdens de NHG-congressen, nascholingsbijeenkomsten en terugkomdagen van huisartsen in opleiding waren de plaatsen waar „kritisch” werd gelezen. Deze bijeenkomsten verliepen over het algemeen erg levendig en maakten veel deelnemers enthousiast voor onderzoek.

De bedoeling van het hieronder afgedrukte verslag is om de lezers van *Huisarts en Wetenschap* kennis te laten maken met het *Draaiboek „kritisch lezen”* en te laten zien hoe dit kan worden gebruikt. Omwille van de leesbaarheid zal een beknopte weergave van de discussie worden gegeven. Het draaiboek zelf is als *bijlage* afgedrukt.

Het is niet eenvoudig geweest om een „kritische lezing” van een reeds verschenen artikel gepubliceerd te krijgen; enige eerdere pogingen zijn gestrand. Wij hebben daarvan geleerd hoe gevoelig deze materie ligt bij auteurs en redacties van de medische vakbladen. Het kritisch lezen (en dat merken we ook in de besprekingen) krijgt erg gauw het image van een afkraak-methode. We willen hier dan ook stellen dat het ons absoluut niet gaat om goed- of afkeuren, maar om het systematisch leren beoor-

delen van onderzoeksverslagen en om het leren wegen van de gepubliceerde resultaten.

Wij zullen daarbij moeten benadrukken dat het niet voldoen aan bepaalde eisen vaak vooral ligt aan praktische onmogelijkheden en dat een onderzoek daarmee niet bij voorbaat waardeloos is. Het doel van de CWO is om onderzoek te stimuleren, niet om het te maken tot een elitaire bezigheid, waarbij we als „aardse” artsen alleen maar fouten zouden kunnen maken. Het kan daarom zin hebben tijdens de bespreking speciaal te letten op positieve categorieën als creativiteit bij de opzet, constructies van meetinstrumenten en dergelijke.

Wij stellen het op prijs dat in deze aflevering van *Huisarts en Wetenschap* een „kritische bespreking” direct na het oorspronkelijke artikel kan worden geplaatst. Wij zijn daarvoor dank verschuldigd aan auteurs en redactiecommissie.

Een onderzoek naar het voorschrijfgedrag van drie huisartsen bij kinderen van 0 - 4 jaar³

De eerste oordelen over het gelezen artikel zijn nogal uiteenlopend. Enkel zijn erg enthousiast, noemen het een schitterend onderzoekje, anderen hebben veel minder affiniteit met de beschreven materie en begrijpen niet goed waarom dit onderzoek is gedaan: wel aardig, maar niet veel bijzonders.

Wat is de vraagstelling van dit onderzoek? Acht u deze voldoende aangegeven? Zijn de begrippen waarmee gewerkt wordt voldoende gedefinieerd? Is er sprake van een operationele definitie?

De vraagstelling is in de inleiding van het artikel duidelijk aangegeven en valt in twee delen uiteen:

1. In hoeverre verschilt het voorschrijfpatroon van drie Nederlandse huisartsen bij 0 - 4 jarige kinderen van het voorschrijfpatroon dat door Grace & Goulds is beschreven?
2. In hoeverre verschillen de drie Nederlandse huisartsen onderling in hun voorschrijfpatroon bij kinderen van 0-4 jaar?

Centraal staat daarin het begrip voorschrijfpatroon; in het paragraafje Methode wordt dit begrip operationeel gemaakt en blijkt het te zijn gedefinieerd als het aantal malen dat volgens de apotheekregistratie een bepaald medicament is afgeleverd, gedeeld door het totaal aantal betrokken kinderen geme-

ten over 5 jaar. Daarbij zijn de medicijnen gerubriceerd naar het voorbeeld van het Engelse onderzoek.

Waarvoor is het relevant dat er een antwoord op de vraagstelling wordt gevonden?

De vraag welke relevantie het onderzoek heeft, levert de CWO-leden duidelijk meer problemen op. In de eerste plaats is door de auteurs niet aangegeven, welk belang zijzelf precies hechten aan hun onderzoek. Het Engelse onderzoek inspireerde hen, maar waarom precies?

Veronderstellingen hebben de lezers wel. Uit de loop van het artikel, en met name uit het hier niet besproken maar wel afgedrukte vervolgartikel, blijkt dat de (vagelijk verwachte) verschillen in voorschrijfgedrag aanleiding kunnen zijn tot discussie over dit aspect van het handelen van de betrokken huisartsen. Een vorm van toetsing dus aan de hand van bijvoorbeeld uitschieters.

Ligt er een hypothese aan de vraagstelling ten grondslag? Waarop is deze gebaseerd?

Een hypothese is niet geformuleerd en lijkt ook niet ten grondslag te liggen aan de vraagstellingen. Wel zou tussen de regels door gelezen kunnen worden dat er een nul-hypothese is, die zou kunnen luiden: „er zijn geen verschillen in voorschrijfgedrag tussen de Engelse en Nederlandse huisartsen en tussen de Nederlandse huisartsen onderling”. Deze hypothese is evenwel theoretisch niet onderbouwd.

Van wat voor soort onderzoek is er sprake? (Onderscheid kan gemaakt worden tussen explorerend en toetsend onderzoek; beide soorten kunnen vervolgens nader getypeerd worden als beschrijvend of als verklarend.)

Als we praten over het soort onderzoek levert ook dit onderdeel van de discussie problemen op. Eensgezind is de groep over het beschrijvende karakter, maar men struikelt over de vraag of hier sprake is van toetsend of van exploratief onderzoek.

De meningen zijn aanvankelijk verdeeld. De opzet van het onderzoek is toetsend: er wordt gekeken of er verschillen bestaan, er wordt vergeleken. Maar omdat die eventuele verschillen niet theoretisch onderbouwd zijn (zie ontbreken hypothese), is het beter te spreken van een explorerend onderzoek.

Welke meetinstrumenten werden hier

gebruikt? (Een meetinstrument kan in dit verband bijvoorbeeld een bloeddrukmeter zijn, maar ook een enquête of de registratie van gegevens over lichamelijk onderzoek of therapie.) Het meetinstrument in het onderzoek is de apotheekregistratie van de aflevering van medicamenten. Alle drie betrokken huisartsen waren apotheekhoudend.

Hoe is het gesteld met de betrouwbaarheid van het meetinstrument? (Betrouwbaarheid kan worden gedefinieerd als de mate waarin twee verschillende onderzoekers onafhankelijk van elkaar tot dezelfde uitslag komen.)

De vraag of het tellen van de gegevens uit deze registratie betrouwbaar (= reproduceerbaar) is, wordt over het algemeen bevestigend beantwoord. Om daarvan zeker te zijn, had men bijvoorbeeld twee onafhankelijke dokters dezelfde gegevens kunnen laten tellen en de uitkomsten vergelijken.

Hoe is het gesteld met de validiteit van de meetinstrumenten? (Validiteit kan worden gezien als de vraag of de onderzoeker meet wat hij wil meten. Men zou kunnen proberen iemands gewicht te meten met een centimeter. De resultaten zullen misschien betrouwbaar zijn – in de zin van steeds dezelfde uitslag – maar niet valide, omdat men iets anders heeft gemeten dan men wilde meten.) De validiteit van dit meetinstrument roept meer vragen op. Meet men met de apotheekregistratie echt het voorschrijfpatroon van de huisarts?

Negatief wordt de validiteit beïnvloed door het feit dat door de assistente ook zonder tussenkomst van de arts medicijnen worden afgeleverd (drogisterijfunctie) en dat medicijnen voorgeschreven door specialisten ook worden meegeteld. Bovendien heeft de samenstelling van de verschillende praktijkpopulaties invloed op het voorschrijven van middelen. Zo zou verondersteld kunnen worden dat de stedelijke bevolking zich eisender opstelt dan de plattelandsbevolking.

Wat kan worden gezegd over de steekproeftrekking? (Hierbij wordt de onderzochte groep onderzoekselementen – patiënten of medicamenten, etc. – gezien als een deel van alle mogelijke onderzoekselementen met datzelfde kenmerk.) **Zijn op grond van de resultaten van dit onderzoek uitspraken te doen over grotere populaties?**

De steekproeftrekking was, zoals de auteurs aangeven, „een steekproef in de tijd”: een steekproef van alle 0 - 4 jari-

gen uit de hele periode die de huisarts als huisarts werkt.

Als we veronderstellen dat in deze periode het voorschrijfgedrag hetzelfde was als in de vroegere en latere loopbaan van de huisarts, kunnen we de gegevens generaliseren voor deze huisarts. Waarschijnlijk is deze steekproeftrekking de beste die de huisartsen hadden kunnen doen.

Hoe vindt u dat de opzet van het onderzoek aansluit op de vraagstelling?

In de vraagstelling wordt gesproken over het nagaan van verschillen tussen de Engelse en Nederlandse huisartsen en de Nederlandse huisartsen onderling. Om dat te doen was het nodig de opzet van het Engelse onderzoek te kopiëren.

Dat is gebeurd. In die zin is de aansluiting van opzet op de vraagstelling goed. Het feit dat door de auteurs gesuggered wordt dat er sprake is van een toetsend onderzoek (significantiebepalingen), doet overtrokken aan.

Hoe worden de resultaten van het onderzoeksproces weergegeven? Wat vindt u van de toepassing van de statistiek?

De resultaten van het onderzoek zijn duidelijk en overzichtelijk weergegeven. Een kleinigheid wordt gemist in de tabellen, zoals het percentage kinderen dat een of ander medicijn kreeg en wat spreidingsmaten. Onduidelijk is welk significantieniveau is gehanteerd.

Welke conclusies worden op grond van de resultaten getrokken? Volgen zij logisch uit de resultaten? Worden zij – in het licht van de representativiteit – van toepassing verklaard op de juiste groep?

De conclusies volgen logisch uit de resultaten en worden terecht niet gegeneraliseerd naar een grotere populatie van huisartsen. Jammer is dat de auteurs in de bespreking als verklaring van verschil tussen Engelse en Nederlandse huisartsen te snel het algemene antibioticabeleid in beide landen opvoeren.

Wat vindt u van de verslaggeving van het onderzoek? Hoe is het gesteld met de samenhang tussen artikel en samenvatting?

Tenslotte vindt men de verslaglegging erg adequaat. Een bondig artikel zonder poespas.

Opgemerkt wordt dat naast het artikel een synopsis staat afgedrukt. Dat is wat anders dan een samenvatting. Voor een samenvatting was volgens sommigen deze weergave van resultaten te uitgebreid.

¹ Eijk JThM van, Gubbels J. Wetenschappelijk onderzoek in de huisartsgeneeskunde. Utrecht: Commissie Wetenschappelijk Onderzoek Nederlands Huisartsen Genootschap, 1983.

² Lamberts H. Wetenschap door de huisarts [Ten geleide]. Huisarts en Wetenschap 1984; 27: 42 - 6.

³ Meyboom WA, Post D, Thomas S. Het voorschrijven van geneesmiddelen aan kleine kinderen. 1. Een onderzoek naar het voorschrijfgedrag van drie huisartsen bij kinderen van 0 - 4 jaar. Huisarts en Wetenschap 1984; 27: 300 - 302.

Bijlage

Draaiboek „kritisch lezen” van de *Commissie Wetenschappelijk Onderzoek van het Nederlands Huisartsen Genootschap, ten behoeve van een bijeenkomst van huisartsen die een artikel (verslag van een wetenschappelijk onderzoek) kritisch willen lezen.*

A. Voorbereiding

1. Ga na waarom de groep die je gaat begeleiden, kritisch wil lezen.

2. Kies een artikel in overleg met de groep en denk hierbij aan:

- relevantie en praktisch nut van het onderzochte onderwerp voor de groep;
- onderzoek gedaan in de eerste lijn;
- aard van het artikel: het artikel moet een verslag zijn van een wetenschappelijk onderzoek (de „oorspronkelijke stukken” in Nederlands Tijdschrift voor Geneeskunde zijn dat vaak), liever geen literatuuroverzicht – of „kookboek”-artikel („bij een verstuikte enkel winde men een zwachtel om de voet”).

3. Stuur het artikel minstens een week voor de bespreking op aan alle deelnemers, vergezeld van de vragen (zie D). Vermeld dat het lezen en maken van de vragen vooraf, voorwaarde is tot zinvolle deelname aan de bespreking.

4. Bereid je voor:

- beantwoord de vragen en lees de basisstof door die betrekking heeft op de moeilijke kanten van dit artikel;
- maak een tijdsindeling voor de vragen;
- als je met z’n tweeën bent, verdeel de taken.

5. Bereid je voor:

6. Beantwoord de vragen en lees de basisstof door die betrekking heeft op de moeilijke kanten van dit artikel;

- maak een tijdsindeling voor de vragen;
- als je met z’n tweeën bent, verdeel de taken.

B. Bijeenkomst

Inleiding

1. Stel je voor (eventueel beide begeleiders):

- lid CWO/NHG;

- geen elite-huisarts of deskundige;

- geïnteresseerd in kritisch lezen: omdat het je tot nu toe..... heeft opgeleverd (noem concrete voorbeelden).

2. Doel „kritisch lezen”: leren beoordelen wat de waarde is van de resultaten van het beschreven onderzoek.

3. „Kritisch lezen” is dus geen afkraakmethode. Het dient ertoe om te leren beoordelen hoe zwaar de conclusies van het onderzoek wegen. Buiten het gezichtsveld blijven vaak de enorme hoeveelheden werk die de onderzoeker verzet heeft en de moeilijkheden die hij heeft ondervonden.

De bespreking

1. Vraag wie het artikel hebben gelezen (beloning gedachte).

2. Als dat niet een duidelijke meerderheid is, besluit of alsnog tijd wordt gemaakt voor lezing (beloning van niet-lezers).

3. Laat alle deelnemers in drie woorden zeggen wat zij van dit artikel vonden. Geef de groep enkele minuten bedenktijd. Noteer de meningen.

4. Loop de vragen langs volgens de tijdsklok.

5. Vat neutraal samen. Laat de groep discussiëren; ga dus niet te gauw in discussie. Laat antwoorden motiveren.

6. Blijf zo nodig buiten de discussie en vervul de rol van uitlegger (van bijvoorbeeld een stuk basisstof), of laat dat over aan de „secondant”.

7. Laat als afsluiting de deelnemers noteren en daarna benoemen wat zij geleerd hebben. Memoreer (eventueel op flap) de oorspronkelijke oordelen over het artikel.

C. Aandachtspunten

1. Ga na hoe je de groep wilt begeleiden; op drie vlakken kun je je deskundig voelen:

- de inhoud: in dit geval dus de basisstof;

- de structuur: het begeleiden van een onderwijsgroep;

- het proces: wat er zich afspeelt op het interpersoonlijke vlak.

2. Als je wilt coördineren, moet je niet zonder duidelijke reden als beoordelaar optreden. Wees je bewust van de rol die je wilt spelen en maak dat duidelijk aan de groep.

3. Geef de tijd om na te denken.

4. Ontstaat een meningsverschil door begripsverwarring? Omschrijf dan zo’n begrip (met elkaar).

5. Doe je niet beter voor dan je bent.

D. Vragen bij het draaiboek

1. Wat is de vraagstelling van dit onderzoek?

Acht u deze voldoende aangegeven? Zijn de begrippen waarmee gewerkt wordt voldoende gedefinieerd? Is er sprake van een operationele definitie?

2. Waarvoor is het relevant dat er een antwoord op de vraagstelling wordt gevonden?

3. Ligt er een hypothese aan de vraagstelling ten grondslag? Waarop is deze gebaseerd?

4. Van wat voor soort onderzoek is er sprake? (Onderscheid kan gemaakt worden tussen explorerend en toetsend onderzoek. Beide soorten kunnen vervolgens nader getypeerd worden als beschrijvend of als verklarend).

5. Welke meetinstrumenten worden hier gebruikt? (Een meetinstrument kan in dit verband bijvoorbeeld een bloeddrukmeter zijn, maar ook een enquête of de registratie van gegevens over lichamelijk onderzoek of therapie.)

6. Hoe is het gesteld met de betrouwbaarheid van de meetinstrumenten? (Betrouwbaarheid kan worden gedefinieerd als de mate waarin twee verschillende onderzoekers onafhankelijk van elkaar tot dezelfde uitslag komen.)

7. Hoe is het gesteld met de validiteit van de meetinstrumenten? (Validiteit kan worden gezien als de vraag of de onderzoeker meet, wat hij wil meten. Men zou kunnen proberen iemands gewicht te meten met een centimeter. De resultaten zullen misschien betrouwbaar zijn – in de zin van steeds dezelfde uitslag – maar niet valide, omdat men iets anders heeft gemeten, dan men wilde meten.)

8. Wat kan gezegd worden over de steekproeftrekking? (Hierbij wordt de onderzochte groep onderzoekselementen – patiënten of medicamenten, etc. – gezien als een deel van alle mogelijke onderzoekselementen met datzelfde kenmerk.) Zijn op grond van de resultaten van dit onderzoek uitspraken te doen over grotere populaties?

9. Hoe vindt U dat de opzet van het onderzoek aansluit op de vraagstelling?

10. Hoe worden de resultaten van het onderzoeksproces weergegeven? Wat vindt U van de toepassing van de beschrijvende statistiek?

11. Wat vindt U van het gebruik van de inductieve statistiek?

12. Welke conclusies worden op grond van de resultaten getrokken? Volgen zij logisch uit de resultaten? Worden zij – in het licht van de representativiteit – van toepassing verklaard op de juiste groep?

13. Wat vindt u van de verslaglegging van het onderzoek? Hoe is het gesteld met de samenhang tussen artikel en samenvatting?