

# Verloskunde gestandaardiseerd

*Susanne Zuidhof-van Arkel MSc, Annemarie Borst-Duineveld, Dr. Cees Lucas*

## Samenvatting

### Doel

Toetsen in hoeverre de kwaliteit van de KNOV-standaarden voldoet aan internationale kwaliteitscriteria en te vergelijken is met de kwaliteit van andere (eerstelijns)beroepsgroepen.

### Methoden

De drie door de KNOV uitgebrachte standaarden worden door twee onderzoekers onafhankelijke beoordeeld middels het AGREE-instrument en vergeleken met vijf CBO-richtlijnen en vijf NHG-standaarden.

### Resultaten

De KNOV-standaarden scoorden op alle domeinen hoger dan 50%, de vijf CBO-richtlijnen scoorden op vier en de vijf NHG-standaarden op drie van de zes domeinen hoger dan 50%. Op vier domeinen scoorden de KNOV-standaarden hoger dan het CBO en op alle domeinen behaalden de KNOV-standaarden een hogere score dan de NHG.

### Conclusie

De kwaliteit van de KNOV-standaarden, beoordeeld door middel van het AGREE-instrument voldoet aan internationale maatstaven (AGREE) en is vergelijkbaar met de kwaliteit van de richtlijnen van het CBO en de NHG-standaarden.

## Inleiding

Binnen de Nederlandse gezondheidszorg wordt, mede als gevolg van het nieuwe verzekeringssysteem en de daaruit voortvloeiende marktwerking, kwaliteit van zorg steeds belangrijker.<sup>[1]</sup> Tijdens de afgelopen jaren heeft in het kader van de kwaliteitszorg de ontwikkeling van richtlijnen voor het medisch handelen, mede onder invloed van Evidence Based Practice, een grote vlucht genomen.<sup>[2]</sup> Het doel van richtlijnen is, naast het ondersteunen van de praktijkvoering het handhaven en

verbeteren van kwaliteit van zorg. Het aantal richtlijnen dat zowel internationaal als nationaal is gepubliceerd, is zo talrijk geworden dat de behoefte aan kwaliteitscriteria en controle steeds groter is geworden.<sup>[3,4]</sup> Ook bij de Koninklijke Nederlandse Organisatie van Verloskundigen (KNOV) staat kwaliteit van zorg hoog in het vaandel. In 2000 publiceerde de KNOV de eerste richtlijn voor het verloskundig handelen. Thans, zeven jaar later zijn er drie richtlijnen ontwikkeld. Als verloskundige beroepsgroep met een relatief korte traditie op het gebied van richtlijnontwikkeling is het goed te weten welk kwaliteitsniveau de richtlijnen hebben ten opzichte van richtlijnen van andere (eerstelijns) beroepsgroepen. Het doel van dit onderzoek is na te gaan in hoeverre de KNOV-standaarden voldoen aan internationaal geaccepteerde kwaliteitscriteria. In dit artikel worden de drie KNOV-standaarden vergeleken met vijf standaarden van het Nederlands Huisartsen Genootschap (NHG) en vijf richtlijnen van het Kwaliteitsinstituut voor de Gezondheidszorg (CBO, voormalig Centraal Begeleidingsorgaan voor de Intercollegiale Toetsing) beoordeeld aan de hand van het Appraisal of Guidelines Research and Evaluation instrument (AGREE).

## Methode

### Richtlijnen

Voor dit onderzoek werd gebruik gemaakt van gegevens uit een eerdere publicatie. Hierin werden vijf NHG-standaarden vergeleken met vijf CBO-richtlijnen, aan de hand van het AGREE-instrument.<sup>[5]</sup> De drie standaarden gepubliceerd door de KNOV worden vergeleken met bovenstaande gegevens. De KNOV standaarden bestaan uit een beschrijving en praktijkkaarten en zijn gebaseerd op een wetenschappelijke onderbouwing. De beschrijving en praktijkkaarten zijn verspreid onder alle leden van de KNOV ([www.knov.nl](http://www.knov.nl)).

### Beoordeling d.m.v. het AGREE-instrument

Het AGREE-instrument bestaat uit 23 items welke worden gescoord op een vierpunts Likert-schaal die zich van zeer eens (4) naar zeer oneens (1) begeeft. De 23 items zijn verdeeld over 6 domeinen: onderwerp en doel; betrokkenheid van belanghebbenden; methodologie; helderheid en presentatie; toepassing en tenslotte onafhankelijkheid van de opstellers. Het doel van het AGREE-instrument is het bieden van een raamwerk om de kwaliteit van klinische richtlijnen te beoordelen. Het instrument is in twee validatieronden getest waarbij 100 richtlijnen uit 11 landen door

*Susanne Zuidhof-van Arkel MSc is als eerstelijns verloskundige verbonden aan Verloskundige Praktijk Hoorn, A Borst-Duineveld is als eerstelijns verloskundige verbonden aan Verloskundige Praktijk Nieuw-Vennep en Dr. C. Lucas is verbonden aan de Afdeling Klinische Epidemiologie, Biostatistiek en Bioinformatica, Universiteit van Amsterdam*

### Correspondentieadres:

*S. Zuidhof-van Arkel,  
Cole Porterhof 135, 1628 TM Hoorn  
0229-216260, [susannezuidhof@kpnplanet.nl](mailto:susannezuidhof@kpnplanet.nl)*

Tabel 1. 1 Geselecteerde richtlijnen

	Publicatiejaar
<b>KNOV-standaarden</b>	
Anemie in de eerstelijns verloskundige praktijk	2000
Hygiëne en infectiepreventie in de eerstelijns verloskundige praktijk (H&I)	2004
Niet-vorderende ontsluiting (NVO)	2006
<b>CBO-richtlijnen</b>	
Osteoporose (tweede herziening)	2002
Seksueel overdraagbare aandoeningen en herpes neonatorum	2002
Chronisch hartfalen (eerste herziening)	2002
Decubitus (tweede herziening)	2002
Behandeling van het mammacarcinoom	2002
<b>NHG-standaarden</b>	
Diagnostiek van mammacarcinoom (eerste herziening)	2002
Reumatoïde artritis (eerste herziening)	2002
Duizeligheid	2002
Anemie	2003
Delier bij ouderen	2003

respectievelijk 194 en 70 beoordelaars zijn beoordeeld. Door meer dan 95 % van de beoordelaars werd het AGREE-instrument zinvol geacht bij het beoordelen van richtlijnen. De betrouwbaarheid was voor de meeste domeinen acceptabel met een Cronbach's alfa 0.64-0.88.<sup>[4]</sup> Het lage aantal items per domein beïnvloedt echter de hoogte van de Cronbach's alfa. De meeste criteria van het AGREE-instrument zijn eerder gebaseerd op theoretische aannamen dan op empirisch bewijsmateriaal. Ze zijn vooral ontwikkeld op basis van discussies tussen onderzoekers van diverse origine. Het AGREE-instrument dient daarom vooral te worden beschouwd als een afspiegeling van de huidige stand van kennis op dit gebied.<sup>[6]</sup>

Om te komen tot een betrouwbaar oordeel werd de beoordeling door 2 beoordelaars (SZvA en ABD) onafhankelijk van elkaar verricht. Vervolgens werd de uitslag ervan vergeleken en bediscussieerd. In geval geen consensus zou worden bereikt, zou het oordeel van de derde beoordelaar (CL) de doorslag geven; dit scenario behoeft echter niet te worden geëffectueerd. Van de itemscore en de domeinscore na consensus werden gemiddelden berekend over de vijf NHG-standaarden, de vijf CBO-richtlijnen en de 3 KNOV-standaarden. Die gemiddelde scores werden vervolgens met elkaar vergeleken. Geen van de beoordelaars was betrokken bij de ontwikkeling van de onderzochte standaarden of richtlijnen.

#### Analyse

Als primaire uitkomstmaat werden bij de beoordeling van de kwaliteit van de standaarden en de richtlijnen de

domeinscores berekend in percentages, zoals beschreven in de instructies bij gebruik van het AGREE-instrument. De ideale statistische maat voor overeenstemming (consensus) tussen beoordelaars is een (gewogen) Kappa, omdat die uitkomstmaat corrigeert voor toeval. De door ons gegenereerde gegevens voldoen echter niet aan de Cohen's criteria. Dit zorgt voor problemen bij de interpretatie, en maakt de Kappa ongeschikt. Om deze reden is er gekozen voor het berekenen van een procentuele overeenstemming. Daartoe werden de scores afkomstig van de beoordeling met het AGREE-instrument eerst gedichotomiseerd in "eens" (scores 4 en 3) en "oneens" (score 2 en 1).

De analyse bestond uit een vergelijking van de gemiddelde domeinscores en itemscores van de KNOV-standaarden, de NHG-standaarden en de CBO-richtlijnen. Voor een variantieanalyse voor de significantie werd een T-test overwogen, echter vanwege het lage aantal 'onderzoekseenheden' moest daarvan worden vanaf gezien. Het berekenen van significantie geeft bij een dergelijk gering aantal KNOV-standaarden mogelijk een vertekend beeld, hetgeen tot over-, of juist onderschatting aanleiding zou kunnen geven.

#### Resultaten

##### Overeenstemming

Het gemiddelde percentage overeenstemming bij het beoordelen van de KNOV-standaarden bedroeg 79.7% voor de consensusbespreking, na inhoudelijke discussie kon volledige overeenstemming worden bereikt.

De oorzaak van verschil in de initiële beoordeling was gelegen in een wisselende interpretatie van relevante paragrafen. De overeenstemming was het laagst bij

item 11, namelijk 33.3%. In 8.7% van de items was het verschil in score 2 punten en bij 1 item (1.4%) bedroeg het een maximaal scoreverschil van 3 punten.

**Tabel 2.** Gemiddelde domeinscores en itemscores van de standaarden en richtlijnen, beoordeeld d.m.v. het AGREE-instrument

	CBO	NHG	KNOV
<b>I Onderwerp en doel</b>	53,3	57,8	74,1
1 Het doel van de richtlijn is specifiek beschreven.	2,0	2,2	3,0
2 De klinische vraag/vragen die in de richtlijn aan de orde komt/komen, is/zijn specifiek beschreven.	3,4	3,0	2,7
3 De patiëntenpopulatie waarop de richtlijn van toepassing is, is specifiek beschreven.	2,4	3,0	4,0
<b>II Betrokkenheid van belanghebbenden</b>	61,7	53,3	58,3
4 De leden van de werkgroep die de richtlijn heeft ontwikkeld, komen uit alle relevante beroepsgroepen.	4,0	2,8	3,3
5 Het perspectief en de voorkeuren van patiënten zijn nagegaan.	3,2	2,0	2,0
6 De beoogde gebruikers van de richtlijn zijn duidelijk benoemd.	3,2	4,0	4,0
7 De richtlijn is getest onder de beoogde gebruikers.	1,0	1,6	1,7
<b>III Methodologie</b>	60,0	33,3	58,7
8 Er zijn systematische methoden gebruikt voor het zoeken naar wetenschappelijk bewijsmateriaal.	3,0	1,0	3,0
9 De criteria voor het selecteren van het wetenschappelijk bewijsmateriaal zijn duidelijk beschreven.	1,8	1,0	2,3
10 De gebruikte methoden om de aanbevelingen op te stellen, zijn duidelijk beschreven.	2,0	1,6	2,0
11 Gezondheidswinst, bijwerkingen en risico's zijn overwogen bij het opstellen van de aanbevelingen.	2,8	3,2	3,7
12 Er bestaat een expliciet verband tussen de aanbevelingen en het onderliggende wetenschappelijk bewijsmateriaal.	3,0	3,0	3,3
13 De richtlijn is voor publicatie door externe experts beoordeeld.	3,0	3,2	3,0
14 Een procedure voor herziening van de richtlijn is vermeld.	4,0	1,0	2,0
<b>IV Helderheid en presentatie</b>	68,3	80,0	94,4
15 De aanbevelingen zijn specifiek en ondubbelzinnig.	3,0	3,6	3,7
16 De verschillende beleidsopties zijn duidelijk vermeld.	3,4	3,4	4,0
17 De kernaanbevelingen zijn gemakkelijk te herkennen.	3,2	3,4	4,0
18 De toepassing van de richtlijn wordt ondersteund met hulpmiddelen.	2,6	3,2	3,7
<b>V Toepassing</b>	44,4	13,3	51,9
19 De mogelijke organisatorische belemmeringen bij het toepassen van de aanbevelingen zijn besproken.	2,4	1,6	3,3
20 De mogelijke kostenimplicaties van het toepassen van de aanbevelingen zijn overwogen.	2,8	1,6	2,3
21 De richtlijn geeft de belangrijkste criteria om na te gaan en te toetsen of de richtlijn wordt gevolgd.	1,8	1,0	2,0
<b>VI Onafhankelijkheid van de opstellers</b>	43,3	0,0	50,0
22 De richtlijn is niet beïnvloed door de opvattingen of belangen van de financierende instantie.	2,4	1,0	3,0
23 Conflicterende belangen van de leden van de werkgroep zijn vastgelegd.	2,2	1,0	2,0

Domeinscores I-VI tonen de gemiddelde percentages berekend volgens de instructies van het AGREE-instrument, van respectievelijk het CBO, de NHG en de KNOV. Itemscores 1-23 betreffen de gemiddelde scores van de richtlijnen / standaarden van desbetreffende organisaties, maximale score 4 (zeer eens), minimale score 1 (zeer oneens).

### Algemeen

Na beoordeling werd een vergelijking gemaakt tussen de domeinscores van de KNOV-standaarden met zowel de CBO-richtlijnen als de NHG-standaarden. Voor deze vergelijking werd gekozen voor een arbitrair afkappunt van 50%, dit om de vergelijking met de resultaten van eerdere publicatie mogelijk te maken.<sup>[5]</sup> De drie KNOV-standaarden scoorden op alle domeinen hoger dan 50%, de vijf CBO-richtlijnen scoorden op vier en de vijf NHG-standaarden op drie van de zes domeinen hoger dan 50%. De KNOV-standaarden scoorden op twee domeinen meer dan 15% hoger dan de CBO richtlijnen ('onderwerp en doel' en 'helderheid en presentatie'). Voor de overige vier domeinen lagen de gemiddelde domeinscores twee maal hoger en twee maal lager dan de scores van de CBO-richtlijnen, de verschillen bedroegen minder dan 10%. De gemiddelde domeinscores van de KNOV-standaarden waren voor alle domeinen hoger dan de gemiddelde domeinscores van de NHG-standaarden. Bij de domeinen 'methodologie', 'toepassing' en 'onafhankelijkheid van de opstellers', scoorden de KNOV-standaarden meer dan 20% hoger dan de NHG-standaarden, de domeinen 'onderwerp en doel' en 'helderheid en presentatie' vertoonden een score van meer dan 10% hoger.

### Onderwerp en doel

Het doel van de KNOV-standaarden Anemie in de eerstelijns verloskundige praktijk (Anemie) en Niet-vorderende ontsluiting (NVO) werd net als in de CBO-richtlijnen en de NHG-standaarden weinig specifiek beschreven. In Hygiëne en infectiepreventie in de eerstelijns verloskundige praktijk (H&I) werd wel een duidelijke beschrijving van het doel gegeven. De klinische vraagstelling was in alle drie de standaarden wel aanwezig maar niet in de eerste oogopslag te vinden. Ten opzichte van de CBO-richtlijnen en NHG-standaarden was bij de KNOV-standaarden de beoogde populatie zeer duidelijk beschreven.

### Betrokkenheid van belanghebbenden

De KNOV-standaarden zijn ontwikkeld voor en door verloskundigen. Dit komt in de standaarden duidelijk naar voren. Daarnaast is ook gebruik gemaakt van de kennis van TNO op het gebied van richtlijnontwikkeling. Het perspectief en de voorkeuren van de patiënten is in

2 van de 3 standaarden niet meegenomen, de standaard NVO heeft hier wel aandacht aan besteed in de vorm van een literatuuronderzoek. In alle drie de standaarden wordt evenals als in de NHG-standaarden melding gemaakt van een commentaarronde. Geen van de standaarden of richtlijnen is vooraf getest onder de gebruikers.

### Methodologie

Zowel in de CBO-richtlijnen, de NHG-standaarden als de KNOV-standaarden werd vermeld dat de literatuur verzameld is door middel van een gestructureerde zoekstrategie in verschillende elektronische databases. Een specifieke beschrijving van de zoekstrategie wordt echter in geen van de gevallen beschreven. In de KNOV-standaarden ontbraken daarnaast ook enkele grote representatieve databases zoals Embase en de Cochrane Library. Een beschrijving van de in- en exclusiecriteria voor het selecteren van de literatuur, de uitkomstmaten, de methodologische kwaliteit en toepasbaarheid van de literatuur werd eveneens niet expliciet vermeld. Ook de methoden voor het opstellen van de aanbevelingen en het tot stand komen van de conclusies werden niet beschreven, wel bestond er een duidelijk verband tussen de aanbevelingen en conclusies en de wetenschappelijke evidentie. Alle standaarden en richtlijnen hebben een externe commentaarronde ondergaan, de gevolgde procedure werd echter niet beschreven. In de CBO-richtlijnen wordt een herzieningsprocedure beschreven, in de NHG- en KNOV-standaarden ontbreekt een dergelijke procesgang. Uitzondering op de bovenstaande gesignaleerde omissie in transparantie vormt de KNOV-standaard H&I, deze scoorde voor het domein methodologie ruim 95% en kenmerkt zich door een duidelijke beschrijving van bovenstaande elementen.

### Helderheid en presentatie

Op dit domein scoorden alle richtlijnen hoog. De aanbevelingen in de standaarden van de KNOV waren evenals als in de standaarden van de NHG redelijk specifiek geformuleerd. Door plaatsing in een kader bleken ze duidelijk herkenbaar en door middel van onder andere stroomdiagrammen was het eenvoudig de verschillende beleidsopties te herkennen. Alle standaarden waren voorzien van praktijkkaarten welke een heldere samenvatting geven van het te volgen beleid.

### Toepassing

In de KNOV-standaarden worden zowel de organisatorische als financiële consequenties die het toepassen van de aanbevelingen vereist benoemd. De diepgang waarmee deze items worden besproken is alleen summier. Dit geldt in meer of mindere mate ook voor de CBO-richtlijnen en NHG-standaarden. Toetsingscriteria, welke het gebruik

van richtlijnen bevorderen, worden, met uitzondering van de CBO-richtlijn Decubitus en KNOV-standaard Anemie, niet genoemd.

#### Onafhankelijkheid van de opstellers

Alle drie de standaarden van de KNOV zijn tot stand gekomen in het kader van het "programma ter bevordering van de doelmatigheid en kwaliteit van zorg door verloskundigen" (PROVER). De standaarden NVO en H&I zijn gefinancierd uit de projectsubsidie "Kwaliteit en Samenwerking 2002-2005" welke door het ministerie van VWS is toegekend. Vertrouwen van de standaarden door belangen of opvattingen van derden is niet te verwachten. Alleen in de KNOV-standaard H&I en de CBO-richtlijn Chronisch hartfalen wordt expliciet vermeld dat de werkgroepleden geen conflicterende belangen hebben.

#### Beschouwing

Volgens de beoordeling met het AGREE-instrument zijn de standaarden van de KNOV kwalitatief goed vergelijkbaar met de standaarden van de NHG en de richtlijnen van het CBO. De overall score van de KNOV lag hoger dan die van de NHG en het CBO. Of het hier een significant verschil betreft, kan aan de hand van dit onderzoek niet geconcludeerd worden, omdat er vanwege het te geringe aantal onderzoekseenheden niet getoetst kon worden op statistische significantie. Onderzoeksgegevens uit een eerdere publicatie<sup>[5]</sup> suggereren dat een verschil in gemiddelde domeinscore van > 20% (n = vijf standaarden) significant is.

Voor de KNOV-standaarden geldt dat evenals voor de NHG-standaarden en CBO-richtlijnen verbetering voornamelijk kan plaats vinden door middel van het inzichtelijk maken van de procedures en methoden. Dit kwam het duidelijkst naar voren bij de beoordeling van de KNOV-standaard H&I. Hierin werd genoemd dat naar aanleiding van de KNOV-standaard anemie in combinatie met de ontwikkeling van kennis op het gebied van richtlijnontwikkeling (in de tussentijd werd het AGREE-instrument gepubliceerd), besloten was tot het toevoegen van een extra hoofdstuk met betrekking tot transparantie van de gebruikte methoden. Dit leidde tot hoge item- en domeinscores. In de daarop volgende standaard bleef dit hoofdstuk echter weer achterwege, hetgeen onmiskenbaar in een lage(re) overall score resulteerde. Hoewel domeinscores nuttig zijn om richtlijnen onderling te vergelijken, blijkt uit de gevoeligheid van het AGREE-instrument voor de wijze van verslaglegging, dat het nauwelijks mogelijk is om objectieve drempelwaarden vast te stellen voor een kwalitatief "goede" of "slechte" richtlijn. Lage scores kunnen zowel veroorzaakt worden door een evident mindere kwaliteit van de standaard,

als ook door een matige verslaglegging. Voor de gebruiker van de richtlijn is vooral de klinische relevantie met betrekking tot de inhoud van belang en juist deze is door middel van het AGREE-instrument niet goed te beoordelen.

#### Conclusie

De kwaliteit van de KNOV-standaarden zoals gemeten met het AGREE-instrument voldoet aan internationale maatstaven en is vergelijkbaar met de kwaliteit van de richtlijnen van het CBO en de NHG-standaarden. Verbetering zou kunnen worden nagestreefd door in de redactie van nieuwe richtlijnen het format van het AGREE-instrument te volgen. ■

#### Literatuur

1. Klazinga N, Lombarts K, Van Everdingen J. Quality management in medical specialties: the use of channels and dikes in improving healthcare in The Netherlands. *Jt Comm J Qual Improv.* 1998;24(5):240-50.
2. Haggard M. The relationship between evidence and guidelines. *Otolaryngol Head Neck Surg.* 2007;137:72-7.
3. Natsch S, Van de Meer JW. The role of clinical guidelines, policies and stewardships. *J Hosp Infect.* 2003;53(3):172-6.
4. The AGREE collaboration. Development and validation of an international appraisal instrument for assessing the Quality of clinical practice guidelines: the AGREE project *Qual Saf Health Care.* 2003;12:18-23.
5. Burgers JS, Boluyt N. Kwaliteit van CBO-richtlijnen en NHG standaarden beoordeeld met het AGREE-Instrument. *Huisarts Wet* 2004;47(9):394-9.
6. The AGREE collaboration. Appraisal of Guidelines for Research & Evaluation (AGREE) Instrument. [www.agreecollaboration.org](http://www.agreecollaboration.org)