

# Evaluatie van het onderdeel medisch handelen van de accreditering

## Vier klinimetrische criteria

Arna van Doorn, Kirsten Kirschner, Margriet Bouma, Jako Burgers, Jozé Braspenning, Richard Grol

### Inleiding

Voor transparantie in de zorg is het nodig om de kwaliteit van zorg te meten met behulp van indicatoren.<sup>1</sup> De huisartsenpraktijk maakt in dit verband gebruik van het Visitatie Instrument Accreditatie (VIA). Verschillende indicatorsets brengen de uiteenlopende aspecten van de huisartsenzorg in kaart.<sup>2</sup> Sinds 2005 zet men in de NHG-Praktijkaccreditering® het VIA in (box 1), dat bestaat uit de onderdelen praktijkorganisatie, patiëntenervaring en medisch handelen. De indicatoren zijn bedoeld om inzicht te krijgen in de kwaliteit van de zorg, waarbij de individuele huisartsenpraktijk het eigen handelen kan vergelijken met andere praktijken (benchmark). Deze wijze van terugkoppeling biedt de huisartsenpraktijk aanknopingspunten ter verbetering van de kwaliteit van de zorg. Het gebruik van dit type informatie neemt steeds meer toe, bijvoorbeeld in het kader van intern kwaliteitsbeleid, multidisciplinaire zorggroepen en publieke informatie. Vragen over de validiteit en betrouwbaarheid van deze informatie zijn dan van eminent belang.

De validiteit en betrouwbaarheid van de onderdelen praktijkorganisatie en patiëntenervaring zijn eerder beschreven.<sup>3,4</sup> Over het

onderdeel medisch handelen is echter nog weinig bekend. Tijdens de ontwikkeling van deze indicatorset is wel voor de afzonderlijke indicatoren de *content validity* vastgesteld – dat wil zeggen dat een panel van experts heeft aangegeven dat het inhoudelijk goede indicatoren zijn om het medisch handelen te evalueren.<sup>2,5</sup> Aan de hand van gegevens die men met het instrument heeft verzameld, is het ook mogelijk andere aspecten van betrouwbaarheid en validiteit te toetsen.<sup>6,7</sup> In dit onderzoek toetsen wij het onderdeel medisch handelen aan de hand van vier klinimetrische criteria.

Ten eerste moeten de diverse indicatorsets van het instrument een breed en divers spectrum aan klinische activiteiten beschrijven die plaatsvinden in de huisartsenpraktijk. Dit betekent dat de indicatorsets onderling niet te zeer met elkaar mogen samenhangen (lage correlatie tussen de indicatorsets). Ten tweede is het van belang dat de indicatoren binnen één set onderling samenhang vertonen, maar niet te zeer overlappende informatie geven (interne consistentie binnen een set). Ten derde willen we dat de indicatorscore een accuraat beeld van de praktijk oplevert. Dit betekent dat we gegevens moeten verzamelen over voldoende patiënten (betrouwbare indicatorscore). Ten vierde moet een

### Samenvatting

Van Doorn AL, Kirschner K, Bouma M, Burgers JS, Braspenning JCC, Grol RPTM. Evaluatie van het onderdeel medisch handelen van de accreditering. Vier klinimetrische criteria. *Huisarts Wet* 2010;53(3):141-6.

**Doel** Het beschrijven van de validiteit en betrouwbaarheid van het onderdeel medisch handelen van het Visitatie Instrument Accreditatie (VIA).

**Methode** Een observationeel onderzoek op basis van gegevens uit het medisch dossier van 82 praktijken.

**Resultaten** De scores op de 7 onderwerpen met betrekking tot chronische aandoeningen, preventieactiviteiten en het antibioticabeleid vertoonden onderling weinig samenhang, waardoor we een redelijk breed beeld kunnen schetsen van het medisch handelen. De scores hadden per onderwerp voldoende samenhang, zodat we ze als 1 concept kunnen beschouwen. Voor het verkrijgen van een betrouwbare indicatorscore zijn gegevens van ten minste 96 patiënten nodig, bij een foutenmarge van 10%. Indien we een steekproef van 40 patiënten nemen, moeten we een foutenmarge van 15% accepteren. Voor het verkrijgen van een betrouwbare

benchmark bleken 233 patiëntenpopulaties (c.q. praktijken) nodig te zijn, bij een foutenmarge van 5%.

**Conclusie** Het onderdeel medisch handelen uit het VIA is geschikt om inzicht te krijgen in het eigen handelen in vergelijking met een landelijke benchmark, mits men rekening houdt met een redelijke foutenmarge. Voor het afleggen van verantwoording of het gebruik in een 'pay-for-performance'-programma willen we de foutenmarge verkleinen. Een steekproef van patiënten volstaat dan niet. We zullen telkens moeten zoeken naar een balans tussen haalbaarheid en rechtvaardigheid.

UMC St. Radboud, IQ healthcare, huispost 114, Postbus 9101, 6500 HB Nijmegen: A.L. van Doorn, junior wetenschappelijk onderzoeker; K. Kirschner, junior wetenschappelijk onderzoeker; dr. J.S. Burgers, huisarts, senior wetenschappelijk onderzoeker; dr. J.C.C. Braspenning, senior wetenschappelijk onderzoeker; prof.dr. R.P.T.M. Grol, hoogleraar Kwaliteit van Zorg, Nederlands Huisartsen Genootschap, Utrecht: dr. M. Bouma, huisarts, wetenschappelijk medewerker.

Correspondentie: A.vanDoorn@iq.umcn.nl

Mogelijke belangenverstreming: niets aangegeven.

### Wat is bekend?

- ▶ Voor transparantie in de zorg is het nodig om de kwaliteit van zorg te meten met een valide en betrouwbaar meetinstrument.

### Wat is nieuw?

- ▶ Het onderdeel medisch handelen van het VIA is een valide en betrouwbaar onderdeel indien men het voor interne kwaliteitsverbetering gebruikt.
- ▶ Bij de interpretatie van de gegevens moet men een acceptabele foutenmarge in acht nemen.
- ▶ Als men verschillende praktijken van elkaar wil kunnen onderscheiden, bijvoorbeeld in het kader van een *pay-for-performance*-systeem, dan zijn er gegevens van grotere aantallen patiënten nodig dan de aantallen die men momenteel includeert.

benchmark op een voldoende aantal praktijken zijn gebaseerd om een stabiel en correct beeld te leveren. Ook moeten de praktijken op grond waarvan men de benchmark berekent representatief zijn voor de praktijk waarvoor men de vergelijking wenst (adequate en betrouwbare benchmark). Voor elk van deze criteria zijn wij nagegaan in welke mate het onderdeel medisch handelen van het Visitatie Instrument Accreditatie (VIA) hieraan voldoet.<sup>2</sup>

## Methode

### Onderzoekspopulatie en instrument

We hebben een observationeel onderzoek gedaan op basis van gegevens van 82 praktijken, die vrijwillig hebben deelgenomen

### Box 1 NHG-Praktijkaccreditering®

Sinds 2005 kunnen huisartsenpraktijken gebruikmaken van de NHG-Praktijkaccreditering®. Huisartsen kunnen gegevens voor de praktijkaccreditering verzamelen met behulp van het Visitatie Instrument Accreditatie (VIA). Het VIA bestaat uit kwaliteitsindicatoren met betrekking tot het medisch handelen, de praktijkorganisatie en de patiëntervaringen, en meet de kwaliteit van (een deel van) het handelen in de huisartsenpraktijk. De gegevens komen uit het medisch dossier en uit vragenlijsten. Daarnaast kan een consulente van NPA bv, de organisatie die de onafhankelijke toetsing uitvoert ten behoeve van het keurmerk NHG-Praktijkaccreditering®, observaties uitvoeren. Aan de hand van de verwerkte gegevens kan de praktijk zichzelf vergelijken met andere praktijken. Op grond van deze vergelijking kan men verbeterplannen opstellen. De praktijk wordt geaccrediteerd op basis van deze verbeterplannen en het resultaat dat de verbeterplannen opleveren, mits aan bepaalde voorwaarden (minimumeisen) is voldaan. De accreditator van NPA toetst dit alles tijdens een audit. De gegevensverzameling met het VIA wordt het eerste jaar uitgevoerd. In de daarop volgende twee jaren vindt opnieuw een audit van de praktijk plaats, waarbij de accreditator toetst of de praktijk de beoogde resultaten heeft behaald, de verbeterplannen voor het komende jaar aan de eisen voldoet en de minimumeisen voor het betreffende jaar in orde zijn. Na drie jaar herhaalt de hele cyclus zich.

aan de NHG-Praktijkaccreditering® in de periode 2005-2006. De praktijken verzamelden gegevens over drie veel voorkomende chronische aandoeningen (diabetes mellitus, astma en COPD), cardiovasculair risicomanagement, enkele specifieke preventieactiviteiten (griepvaccinatie en cervixscreening) en het antibioticabeleid. Naast procesindicatoren kwamen ook structuurindicatoren en uitkomstindicatoren voor.<sup>8</sup> We hebben ons voor de toetsing van de criteria vooral gericht op de procesindicatoren, omdat deze de meest betrouwbare informatie geven over de kwaliteit van het eigen handelen. In *tabel 1* staan als voorbeeld de negen gebruikte procesindicatoren in het kader van diabetes. Voor cardiovasculair risicomanagement gebruikten we acht procesindicatoren, voor COPD en astma vijf respectievelijk vier indicatoren, en voor griepvaccinatie, cervixscreening en antibioticabeleid ieder één indicator. Voor een aantal berekeningen hebben we ook enkele uitkomstindicatoren meegenomen, om ook zicht te krijgen op dit type informatie (met betrekking tot diabetes drie en CVR twee indicatoren). Een overzicht van alle indicatoren die we in het onderzoek hebben meegenomen staat in *tabel 6* (pag 146).

### Gegevensverzameling

De huisartsenpraktijk verzamelde zelf de benodigde gegevens en het NPA/NHG stelde deze aan de onderzoekers ter beschikking. De bron van de gegevens was het medisch dossier. De gegevens waren echter meestal niet gemakkelijk uit het Huisartsen Informatie Systeem te extraheren, enerzijds omdat daartoe software ontbrak en anderzijds omdat de gegevens niet op uniforme wijze waren geregistreerd. Daarom konden praktijken er ook voor kiezen om gegevens over een steekproef van veertig patiënten te verzamelen. Als het totale aantal patiënten met de betreffende aandoening niet groter was dan veertig patiënten, moest men de gegevens van alle patiënten verzamelen. Het VIA vraagt gegevens te verzamelen voor dat deel van de betreffende patiëntengroep dat door de huisarts is behandeld. In plaats daarvan rapporteerde men echter vaak over alle patiënten met de betreffende aandoening. Voor de benodigde gegevens over het antibioticabeleid adviseerden we een beroep te doen op de preferente apotheker. De gegevensverzameling kon betrekking hebben op 1 huisarts of meer huisartsen als deze de patiëntenpopulatie deelden. Van de 82 praktijken includeerden we in totaal 97 patiëntenpopulaties; van elke populatie hadden we gegevens over alle onderwerpen.

### Analyse

Om te achterhalen of de praktijken uit ons onderzoek representatief waren voor alle praktijken in Nederland hebben we onze onderzoekspopulatie vergeleken met de Nederlandse huisartsenpraktijken wat betreft praktijktype, urbanisatiegraad en apotheekhoudendheid. Bovendien keken we naar het aantal patiënten per fte huisarts. Om te toetsen of de onderwerpen binnen het medisch handelen uit het VIA voldoende van elkaar verschilden, hebben we de mate van samenhang tussen de verschillende onderwerpen (criterium 1) bepaald met behulp van Pearson's correlatiecoëfficiënt. Voor het beoordelen van de interne consistentie (criterium 2) heb-

**Tabel 1** Procesindicatoren in het kader van diabetes

Indicator	Omschrijving
1	Percentage diabetespatiënten met een volledig ingevuld risicoprofiel
2	Percentage diabetespatiënten bij wie in het afgelopen jaar minstens drie maal een glucosemeting is gedaan
3	Percentage diabetespatiënten bij wie in het afgelopen jaar het HbA1c is bepaald
4	Percentage diabetespatiënten bij wie in het afgelopen jaar de bloeddruk is bepaald
5	Percentage diabetespatiënten bij wie in het afgelopen jaar het cholesterolgehalte is bepaald
6	Percentage diabetespatiënten bij wie in het afgelopen jaar het creatininegehalte is berekend
7	Percentage diabetespatiënten bij wie in de afgelopen twee jaar een oogonderzoek heeft plaatsgevonden
8	Percentage diabetespatiënten bij wie in het afgelopen jaar een voetonderzoek heeft plaatsgevonden
9	Percentage diabetespatiënten dat een cholesterolverlagend medicament gebruikt

ben we de samenhang tussen de verschillende indicatoren van 1 onderwerp bepaald met behulp van Cronbach's  $\alpha$ .<sup>9</sup> Een Cronbach's  $\alpha$  tussen de 0,7 en 0,8 is gewenst, maar een  $\alpha$  boven de 0,6 acht men voor dit soort indicatoren ook acceptabel.<sup>10</sup> Per indicator hebben we voor elke patiëntenpopulatie het aantal benodigde patiënten berekend om tot een betrouwbare score te komen (criterium 3). Deze berekening baseerden we op de gevonden indicatorscore. We hebben een betrouwbaarheidsniveau van 95% aangehouden en een foutenmarge van zowel 5% als 10% getoetst. Per indicator hebben we vervolgens het maximum, de 75e percentielscore, het gemiddelde en de standaarddeviatie van de gevonden aantallen berekend om een goed beeld te krijgen van de scores en de mate van spreiding. Voor een schatting van het aantal benodigde patiëntenpopulaties voor een betrouwbare benchmark (criterium 4) maakten we ook gebruik van een powerberekening.<sup>11</sup> Hiervoor hebben we een betrouwbaarheidsniveau van 95% aangehouden en toetsten we wederom een toegestane foutenmarge van 5 en 10%.

## Resultaten

### Onderzoekspopulatie

De meeste kenmerken van de huisartsenpraktijken uit ons onderzoek komen redelijk overeen met de kenmerken van de Nederlandse huisartsenpraktijken (tabel 2). Wel lijken er relatief minder solopraktijken voor te komen in ons onderzoek. Verder is het aantal patiënten per fte huisarts van onze onderzoekspopulatie hoger dan het landelijk aantal inwoners per fte huisarts.

### Criterium 1: geringe onderlinge samenhang indicatorsets

We vonden geen sterke correlaties tussen de verschillende onderwerpen die aan bod kwamen binnen het onderdeel medisch handelen (zie tabel 3). Het onderwerp antibiotica hing met geen enkel ander onderwerp samen. Cervixscreening bleek enigszins samen te hangen met cardiovasculair risicomanagement; deze correlatie was echter zwak. Van de overige vijf onderwerpen hing astma het sterkst samen met de andere

onderwerpen (diabetes, COPD, cardiovasculair risicomanagement en griepvaccinatie). Verder stelden we enige samenhang vast tussen diabetes en zowel cardiovasculair risicomanagement als griepvaccinatie.

### Criterium 2: interne consistentie van de indicatorsets

In tabel 4 staan de gemiddelde scores op de indicatoren per onderwerp, de standaarddeviatie en de samenhang van de indicatoren per onderwerp (Cronbach's  $\alpha$ ). De negen procesindicatoren die op diabetes betrekking hebben bleken onderling de grootste samenhang te vertonen. Deze indicatoren hadden een Cronbach's  $\alpha$  van 0,73. Deze score kan variëren tussen de 0 en de 1, waarbij 0 staat voor totaal geen samenhang en 1 voor sterke samenhang. Als de score richting de 0,9 of hoger gaat, is de samenhang zo sterk dat de indicatoren overlappende informatie geven. Men kan er dan voor kiezen indicatoren weg te laten. Scores richting de 0 duiden erop dat de indicatoren verschillende onderwerpen betreffen, wat betekent dat men de indicatoren beter niet kan samennemen. De 8 procesindicatoren met betrekking tot cardiovasculair risicomanagement hingen redelijk sterk samen, evenals de 5 indicatoren betreffende COPD (respectievelijk  $\alpha = 0,64$  en  $\alpha = 0,67$ ). De 4 astma-indicatoren hadden een iets minder sterke samenhang ( $\alpha = 0,56$ ). Een beschrijving van de samenhang aan de hand van Cronbach's  $\alpha$  is alleen mogelijk voor onderwerpen met meer dan 1 indicator.

**Tabel 2** Praktijkenmerken van de onderzoekspopulatie in vergelijking met alle Nederlandse praktijken (2008)

Praktijkvorm	Onderzoekspopulatie NHG-Praktijkaccreditering® n = 82 praktijken		Alle Nederlandse huisartsenpraktijken* n = 4235 praktijken
	n	%	%
Solopraktijk	18	22,0	42,3
Duopraktijk	18	22,0	31,5
Groepspraktijk of gezondheidscentrum	24	29,3	26,1
HOED	15	18,3	-
Anders	7	8,5	-
Urbanisatiegraad†			
Zeer sterk/sterk stedelijk	33	40,2	46,7
Matig/weinig stedelijk	38	46,3	41,1
Niet stedelijk	11	13,4	12,2
Apotheekhoudend			
Ja	3	3,7	7,3
Nee	79	96,3	92,7
Aantal patiënten per fte huisarts	2444		
Aantal inwoners per fte huisarts			2322

\* Gegevens van het NIVEL, 1-1-2008<sup>12</sup>

† Zeer sterk/sterk stedelijk  $\geq 1500$  adressen per km<sup>2</sup>; matig/weinig stedelijk = 500-1500 adressen per km<sup>2</sup>; niet stedelijk  $\leq 500$  adressen per km<sup>2</sup>

HOED: Huisartsen Onder Een Dak

**Tabel 3** Samenhang onderwerpen medisch handelen, uitgedrukt in Pearson's correlatiecoëfficiënt over de gemiddelden

	Diabetes	COPD	Astma	CVR	Griepvaccinatie	Cervixscreening	Antibiotica
Diabetes	X	0,15	0,31*	0,43*	0,29*	0,19	-0,10
COPD		X	0,40*	0,09	0,20	-0,13	0,12
Astma			X	0,48*	0,30*	0,16	0,06
CVR				X	0,08	0,27*	0,19
Griepvaccinatie					X	0,17	0,17
Cervixscreening						X	-0,07
Antibiotica							X

\* Correlatie is significant,  $p < 0,01$   
 CVR: cardiovasculair risicomanagement

**Tabel 4** Samenhang van de indicatoren per onderwerp, uitgedrukt in Cronbach's  $\alpha$  (n = 97)

Onderwerp	Aantal indicatoren	Cronbach's $\alpha$	Gemiddelden in %	SD in %
Diabetes	9	0,73	72,4	11,2
COPD	5	0,67	63,1	15,4
Astma	4	0,56	49,3	17,1
CVR	8	0,64	49,6	11,4
Griepvaccinatie	1	-	85,6	8,4
Cervixscreening	1	-	67,6	16,5
Antibiotica	1	-	14,3	11,5

SD: standaarddeviatie; CVR: cardiovasculair risicomanagement

**Criterium 3: betrouwbare indicatorscores per patiëntenpopulatie**

Gegeven de gevonden indicatorscores berekenden we voor elke patiëntenpopulatie per indicator het benodigde aantal patiënten. In *tabel 5* staat per onderwerp het minimaal aantal benodigde patiënten. Voor onderwerpen met meer indicatoren hebben we dit voor elke indicator berekend; *tabel 5* geeft het laagste en hoogste aantal benodigde patiënten per onderwerp weer. Voor de 9 diabetesprocesindicatoren was het laagste aantal 363 en het hoogste aantal 384 patiënten voor de berekening van een betrouwbare score bij een foutenmarge van 5%, gegeven de behaalde indicatorscores. Voor bijna alle onderwerpen zien we dat er minimaal 384 patiënten nodig zijn. Als we een afwijking van 10% acceptabel vinden, dan zouden we 96 patiënten moeten beoordelen om betrouwbare scores te kunnen berekenen. De gemiddelde populatiegrootte van de deelnemende praktijken was 3917 patiënten. Bij deze praktijkgrootte kunnen we betrouwbare indicatorscores (met een afwijking van 10%) alleen leveren voor aandoeningen met een minimale prevalentie van 2,4%. Dit geldt voor diabetes, astma en cardiovasculair risicomanagement (RIVM-site). De prevalentie van COPD ligt net iets onder de 2,4%. In huisartsenpraktijken is mogelijk sprake van onderregistratie, wat betekent dat we bij een verbeterde registratie ook voor COPD voldoende patiënten kunnen

includeren om betrouwbare indicatorscores te kunnen leveren.

**Criterium 4: betrouwbare benchmark**

Het aantal benodigde patiëntenpopulaties ten behoeve van de berekening van een betrouwbare benchmark hangt af van de spreiding tussen de verschillende scores op de betreffende indica-

tor. De standaarddeviatie per indicator varieerde tussen 5,1 bij de diabetesuitkomstindicator 'percentage diabetespatiënten met een HbA1c boven de 8,5' en 39,0 bij de diabetesprocesindicator 'percentage patiënten met een volledig ingevuld risicoprofiel'. Omdat deze spreiding tussen de verschillende indicatoren sterk varieerde zijn er ook grote verschillen in het aantal benodigde patiëntenpopulaties per indicator. Zo varieert het minimum aantal benodigde patiëntenpopulaties voor de diabetesindicatoren tussen de 8 en 233 bij een foutenmarge van 5% (zie *tabel 5*). Om voor alle indicatoren een betrouwbare benchmark te kunnen berekenen zijn 233 patiëntenpopulaties nodig bij een foutenmarge van 5% en 58 populaties bij een foutenmarge van 10%. De huidige benchmarkcijfers die men ter referentie bij het VIA aanbiedt, zijn gebaseerd op 259 populaties, wat voor deze onderzoekspopulatie ruim voldoende is om een maximale foutenmarge van 5% te handhaven.

**Beschouwing**

Uit de resultaten blijkt dat het onderdeel medisch handelen van het VIA een relatief breed en divers beeld schetst van de huisartsenzorg. De onderwerpen die hierin zijn meegenomen, hangen onderling weinig samen. De samenhang die we wel vonden valt te verklaren doordat er overlap is tussen de verschillende groe-

**Tabel 5** Aantal benodigde patiënten en aantal benodigde patiëntenpopulaties voor benchmark, hoogste en laagste aantal per onderwerp

Onderwerp (aantal)	Aantal benodigde patiënten per indicator		Aantal benodigde populaties benchmark	
	Foutenmarge 5%	Foutenmarge 10%	Foutenmarge 5%	Foutenmarge 10%
Diabetes (p, 9)	363-384	91-96	8-233	2-58
Diabetes (u, 3)	272-384	68-96	4-35	1-9
COPD (5)	384-384	96-96	46-146	11-37
Astma (4)	384-384	96-96	64-175	16-44
CVR (p, 8)	384-384	96-96	28-161	7-40
CVR (u, 2)	384-384	96-96	39-46	10-11
Griepvaccinatie (1)	372	93	11	3
Cervixscreening (1)	384	96	42	10
Antibiotica (1)	383	96	20	5

p = procesindicatoren, u = uitkomstindicatoren  
 CVR: cardiovasculair risicomanagement

pen. Zo kan een patiënt die is geïnccludeerd bij de indicatoren met betrekking tot cardiovasculair risicomangement ook diabetes hebben. Deze patiënt komt dan ook in aanmerking voor een griepvaccinatie. In de toekomst kan men meer onderwerpen aan het instrument toevoegen, zodat de evaluatie completer wordt en het veld van de huisartsgeneeskunde breder wordt gedekt. De onderwerpen die in de huidige versie aan bod komen toetst men aan de hand van een set van indicatoren die onderling redelijk samenhangen. Alleen Cronbach's  $\alpha$  betreffende astma was aan de lage kant. Een mogelijke verklaring voor een lagere samenhang van astma-indicatoren is dat men voor astma een minder strikt controlebeleid hanteert, vooral als de patiënt geen of weinig medicatie gebruikt. Voor mensen met diabetes, cardiovasculair risicomangement en COPD, waarbij de huisarts ook de gemeenten indicatoren moet uitvoeren, is het controlebeleid strikter. Een redelijke tot goede samenhang tussen indicatoren duidt erop dat we één achterliggend concept meten. Ten aanzien van het verkrijgen van betrouwbare scores komt er bij de verschillende indicatoren vaak eenzelfde maximum aantal benodigde patiënten naar voren. Dit komt doordat er bij bijna alle indicatoren minimaal één patiëntenpopulatie een score van 50% had; statistisch gezien is bij deze score het grootste aantal patiënten nodig voor een betrouwbare score. Hoe meer de scores de 0% of 100% naderen, des te kleiner het aantal benodigde patiënten wordt. De betrouwbaarheid van de scores hangt verder ook af van de betrouwbaarheid van de gegevens op grond waarvan men deze scores heeft berekend. Indien de gegevens niet betrouwbaar zijn verzameld, bijvoorbeeld door problemen bij de dataverzameling, zullen de scores ook nooit betrouwbaar zijn. Uniforme rapportage is daarom ook van het grootste belang. Bovendien moet men van tevoren bepalen welke gegevens men in- en excludeert.

De vraag hoeveel patiënten nodig zijn voor een betrouwbare score en welk betrouwbaarheidsinterval men nog acceptabel acht, hangt sterk af van het doel van de meting. Voor interne kwaliteitsverbetering hoeft de score niet op de procent nauwkeurig te zijn om een goed beeld te geven van mogelijke verbeterpunten. Waar we de grens precies moeten leggen is arbitrair en wordt mede bepaald door de haalbaarheid. Een praktijk heeft slechts een beperkt aantal patiënten uit de doelgroep in de praktijk. Ook is er een grens wat betreft de tijdsinvestering door de huisarts. De steekproef van 40 patiënten die momenteel bij het VIA is toegestaan, is betrouwbaar zolang men een foutenmarge van 15% accepteert. Bij gebrek aan grotere aantallen gegevens kunnen deze scores wel een redelijke indicatie bieden voor interne kwaliteitsverbetering. Bij het gebruik van het instrument stelt men strengere eisen aan het instrument naarmate dat beter moet kunnen differentiëren tussen verschillende praktijken of scores nauwkeuriger moet weergeven. Indien men het instrument wil gebruiken om praktijken van elkaar te onderscheiden, bijvoorbeeld in het kader van een *pay-for-performance*-systeem, zal men grotere aantallen patiënten moeten includeren. Voor dit doel is een foutenmarge van 5% of minder gewenst, wat betekent dat men alleen onderscheid kan maken op zorggroepniveau of tussen grotere praktijken.

## Abstract

Van Doorn AL, Kirschner K, Bouma M, Burgers JS, Braspenning JCC, Grol RPTM. Evaluation of clinical indicators. Four reliability and validity issues. *Huisarts Wet* 2010;53(3):141-6.

**Aim** To describe four reliability and validity issues regarding the clinical indicators from the Visitation Instrument Accreditation (VIA). Based on this information practices needed to start improvement projects in order to get accreditation.

**Method** An observational study based on the medical records of 82 practices.

**Results** The indicators that covered chronic disease management (diabetes, COPD, asthma and cardiovascular risk management), prevention activities (influenza vaccination, cervical cancer screening) and antibiotics policy were correlated weakly, suggesting that the instrument provided a rather broad scope of the practice when it comes to chronic disease management and prevention. Furthermore, the different topics were each measured by indicators that had a sufficient coherence, which suggested that they measured a clear underlying concept. To achieve a reliable indicator score, data from at least 96 patients were necessary when 10% error is allowed. VIA allows to take a sample of 40 patients, but in that case the error margin increases to 15%. To establish a reliable benchmark we needed 233 practices when 5% error is allowed.

**Conclusion** The clinical indicators from VIA are reliable and valid and can be used by a general practice to gain insight into their own performance compared to others. For practice policy on quality improvement an error margin of 10-15% around the indicator score on practice level seems to be acceptable. In case of accountability or a pay-for-performance program we would be more comfortable with a smaller error margin. A sample of 40 patients wouldn't do. However, it is again and again a search to find the balance between feasibility and justice.

In ons onderzoek hebben we de berekeningen niet uitgevoerd op praktijkniveau, maar op patiëntenpopulatie niveau. In de meeste gevallen was er maar één populatie in de praktijk. Voor praktijken die gegevens aanleverden voor meer, gescheiden populaties kunnen we deze samennemen om een beter beeld te geven van de gehele praktijk. Bij de interpretatie van de gegevens moet men bedenken dat de scheiding tussen 'betrouwbaar' en 'onbetrouwbaar' niet zo hard is als deze wellicht lijkt. Ook indien men op basis van het aantal geïnccludeerde patiënten moet concluderen dat een indicatorscore niet betrouwbaar is, kan de indicatorscore toch nog een redelijk beeld van de praktijk geven. Het is echter niet 95% zeker dat de score representatief is, wat wil zeggen dat men voorzichtig moet zijn bij het trekken van conclusies op basis van de gevonden cijfers. Door de foutenmarge te verhogen naar 15%, daalt het benodigde aantal patiënten of patiëntenpopulaties. Al naar gelang het doel kan men variëren met de gekozen foutenmarge en het aantal geïnccludeerde patiënten om een zo betrouwbaar mogelijke score te krijgen.

## Conclusie

Uit ons onderzoek blijkt dat het onderdeel medisch handelen van het VIA wat betreft de vier onderzochte criteria een betrouwbaar en valide onderdeel is. Hierbij geldt dat men bij de interpretatie van gegevens een acceptabele foutenmarge in acht moet nemen. Uniforme rapportage is van groot belang om de betrouwbaarheid van de gegevensverzameling te verbeteren. Voor de huisarts is het een waardevol instrument om inzicht te krijgen in het eigen handelen in vergelijking met landelijke referentiecijfers. Op basis van dit instrument kunnen huisartsen plannen ter verbetering van de kwaliteit van het medisch handelen opstellen. Met behulp van het instrument kan men ook landelijk kijken naar de kwaliteit van de zorg van het medisch handelen. Het aantal benodigde patiënten over wie men gegevens moet verzamelen, zal vanwege de hogere eisen aan de betrouwbaarheid wel groter moeten zijn, dan wanneer men het instrument voor eigen gebruik inzet.

## Financiering

Het Nederlands Huisartsen Genootschap heeft dit onderzoek gefinancierd in het kader van het VIA-project.

## Literatuur

1 Grol R, Wensing M. Implementatie; effectieve verbetering van de patiëntenzorg. 3e druk. Maarssen: Elsevier gezondheidszorg, 2006.

- 2 Braspenning J, Dijkstra R, Tacken M, Bouma M, Witmer H. Visitation instrument accrediting (VIA®). Nijmegen/Utrecht: WOK/NHG/NPA, 2007.
- 3 Grol R, Wensing M, Mainz J, Ferreira P, Hearnshaw H, Hjortdahl P, et al. Patients' priorities with respect to general practice care: an international comparison. European Task Force on Patient Evaluations of General Practice (EUROPEP). *Fam Pract* 1999;16:4-11.
- 4 Van den Hombergh P, Grol R, Van den Hoogen HJ, Van den Bosch WJ. Assessment of management in general practice: validation of a practice visit method. *Br J Gen Pract* 1998;48:1743-50.
- 5 Braspenning JCC, Pijnenborg L, In 't Veld CJ, Grol RPTM, redactie. Werken aan kwaliteit in de huisartsenpraktijk. Indicatoren gebaseerd op de NHG-Standaarden. Houten: Bohn Stafleu van Loghum, 2005.
- 6 Campbell S, Braspenning J, Hutchinson A, Marshall M. Research methods used in developing and applying quality indicators in primary care. *Qual Saf Health Care* 2002;11:358-64.
- 7 Streiner DL, Norman GR. Health measurement scales: a practical guide to their development and use. 3e druk. Oxford: Oxford University Press, 2003.
- 8 Donabedian A. Explorations in quality assessment and monitoring. The definition of quality and approaches to its assessment. Michigan: Health Administration Press, 1980.
- 9 Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika* 1951;16:297-334.
- 10 Bland JM, Altman DG. Statistics notes: Cronbach's alpha. *BMJ* 1997;314:572.
- 11 Moore DS, McCabe GP. Statistiek in de praktijk; Schoonhoven: Academic Service, 1999.
- 12 Hingstman L, Kenens RJ. Cijfers uit de registratie van huisartsen – peiling 2008. Utrecht: NIVEL, 2009.

Tabel 6 Overzicht van alle indicatoren

Onderwerp	Nummer	Omschrijving
Diabetes procesindicatoren	1	Percentage diabetespatiënten met een volledig ingevuld risicoprofiel
	2	Percentage diabetespatiënten bij wie in het afgelopen jaar minstens drie maal een glucosemeting is gedaan
	3	Percentage diabetespatiënten bij wie in het afgelopen jaar het HbA1c is bepaald
	4	Percentage diabetespatiënten bij wie in het afgelopen jaar de bloeddruk is bepaald
	5	Percentage diabetespatiënten bij wie in het afgelopen jaar het cholesterolgehalte is bepaald
	6	Percentage diabetespatiënten bij wie in het afgelopen jaar het creatininegehalte is berekend
	7	Percentage diabetespatiënten bij wie in de afgelopen twee jaar een oogonderzoek heeft plaatsgevonden
	8	Percentage diabetespatiënten bij wie in het afgelopen jaar een voetonderzoek heeft plaatsgevonden
uitkomstindicatoren	9	Percentage diabetespatiënten dat een cholesterolverlagend medicament gebruikt
	1	Percentage diabetespatiënten met een HbA1c boven de 8,5
	2	Percentage diabetespatiënten met een bloeddruk binnen de normaalwaarden (systolisch onder de 150 mmHg en diastolisch onder de 85 mmHg)
COPD	3	Percentage diabetespatiënten met een cholesterolwaarde lager dan 5,0 mmol/l
	1	Percentage patiënten met COPD met bevestiging van diagnose middels spirometrie
	2	Percentage patiënten met COPD met een spirometriemeting in het afgelopen jaar
	3	Percentage patiënten met COPD van wie het rookgedrag bekend is
	4	Percentage rokende patiënten met COPD met een stoppen met roken advies in het afgelopen jaar
5	Percentage patiënten met wie in het afgelopen jaar contact geweest is vanwege COPD	
Astma	1	Percentage patiënten met astma met bevestiging van de diagnose middels piekstroommeter of spirometrie
	2	Percentage patiënten met astma (> 12 jaar) van wie het rookgedrag bekend is
	3	Percentage rokende patiënten met astma met een stoppen met roken advies in het afgelopen jaar
	4	Percentage patiënten met wie in het afgelopen jaar contact geweest is vanwege astma
CVR procesindicatoren	1	Percentage patiënten met hoogrisico voor CVR van wie het rookgedrag bekend is
	2	Percentage rokende CVR hoogrisicopatiënten met een stoppen met roken advies in het afgelopen jaar
	3	Percentage CVR hoogrisicopatiënten met gecontroleerde bloedsuiker middels glucosemeting in het afgelopen jaar
	4	Percentage CVR hoogrisicopatiënten met gecontroleerde bloeddruk in het afgelopen jaar
	5	Percentage CVR hoogrisicopatiënten met een cholesterolbepaling of cholesterol/HDL-ratio in het afgelopen jaar
	6	Percentage CVR hoogrisicopatiënten met een volledig ingevuld risicoprofiel
	7	Percentage CVR hoogrisicopatiënten dat cholesterolverlagers gebruikt
	8	Percentage patiënten met CVR in de anamnese dat antistollingsmiddelen gebruikt
uitkomstindicatoren	1	Percentage CVR hoogrisicopatiënten met een bloeddruk binnen de normaalwaarden (systolisch onder de 160 mmHg en diastolisch onder de 90 mmHg)
	2	Percentage CVR hoogrisicopatiënten met cholesterolverlagers en een cholesterolwaarde lager dan 5,0 mmol/l
Griepvaccinatie	1	Percentage gevaccineerde hoogrisicopatiënten in de praktijk of percentage gevaccineerde patiënten van 65 jaar en ouder*
Cervixscreening	1	Percentage vrouwen uit het doelcohort van het bevolkingsonderzoek waarbij een uitstrijk is gedaan
Antibiotica	1	Percentage smalspectrumantibiotica ten opzichte van alle voorgeschreven antibioticamiddelen voorgeschreven in het afgelopen jaar

\* Aangezien praktijken vaak niet in staat waren beide indicatoren in te vullen, hebben we er voor gekozen om als indicatorscore voor dit onderwerp de hoogste waarde van deze twee indicatoren te gebruiken.