

Tussen gouden standaard en gouden kalf

Functie en toepassing van NHG-standaarden

NHG-standaarden hebben een belangrijke maatschappelijke functie, omdat ze concreet aangeven hoe de beroepsgroep werk maakt van haar kwaliteitsbeleid. Intern – binnen de beroepsgroep – zijn NHG-standaarden bedoeld als referentiekader, zowel voor individuele huisartsen, als voor toetsgroepen. Daarnaast is geleidelijk een tweede toepassing opgekomen: de NHG-standaard als maat voor onderzoek naar goed klinisch handelen van de huisarts. Bij de eerste toepassing – de huisarts toetst zelf zijn handelen, al of niet samen met collega's – gaat het om het handhaven of verbeteren van de kwaliteit van handelen. De NHG-standaard is dan een van de middelen die toetsgroepen ter beschikking staan. Bij de tweede toepassing wordt de kwaliteit van handelen gemeten door een ander. Een favoriete werkwijze bij deze toepassing is het trekken van een representatieve steekproef uit alle Nederlandse huisartsen. Bij die huisartsen wordt dan op een of andere manier vastgesteld of ze een bepaalde NHG-standaard volgen. Vervolgens worden conclusies getrokken over de kwaliteit van het handelen van die gemiddelde huisarts.¹⁻³ Daarmee krijgen de NHG-standaarden – bedoeld of onbedoeld – gemakkelijk het aura van een gouden standaard. 'Doet de huisarts wat hij zou moeten doen?', is dan de vraagstelling.

In het kielzog van deze ontwikkeling lijken de gebruikte methoden van gegevensverzameling soms gekozen op grond van de toegankelijkheid van die gegevens in plaats van op grond van een interessante vraagstelling. Dat leidt tot kaartenbak-onderzoek, enquêtes en secundaire analyses van bestaande morbiditeitsregistraties.

Deze werkwijze, in combinatie met de achterliggende opvatting over NHG-standaarden, was voor ons aanleiding de vraag op te werpen of een NHG-standaard inderdaad beschouwd mag worden als een soort *gouden* standaard voor kwaliteit van handelen. Daarnaast vroegen we ons af wat passende gegevens zouden zijn om de kwaliteit van zorg aan af te meten.

Individu of populatie graadmeter?

Eddy onderscheidt drie soorten referentiekaders voor professioneel handelen: opties, richtlijnen en standaarden.⁴ Opties zijn het meest flexibel; het zijn keuzemogelijkheden: ze kunnen al of niet worden gevolgd. Richtlijnen zijn wat minder flexibel: het zijn adviezen met een normerend karakter, maar er kan beargumenteerd van worden afgeweken. Standaarden zijn niet flexibel. Standaarden zijn voorschriften waarvan niet afgeweken mag worden. Wie wel afwijkt, doet het niet goed.

NHG-standaarden zijn dus geen standaarden, maar richtlijnen, zoals blijkt uit de vaste kaders bij elke standaard over inbreng van de patiënt en het persoonlijk inzicht van de huisarts.^{5,6} Ze beschrijven wat in de meeste gevallen verstandig handelen is, zodat naar de verwachting van de

Van Duijn NP, Meyboom-de Jong B. Tussen gouden standaard en gouden kalf. Functie en toepassing van NHG-standaarden [Commentaar]. Huisarts Wet 1995; 38(1): 3-6.

beroepsgroep de meeste patiënten er voordeel, of ten minste geen nadeel, van zullen hebben.

De voordelen van richtlijnen kunnen voordelen voor individuele patiënten zijn of voordelen voor populaties. Maar vaak gaat het om een compromis dat door middel van consensus wordt vastgesteld. Een voorbeeld van zo'n compromis is het terughoudende antibioticabeleid bij keelpijn, waarbij neveneffecten van antibiotica op populatieniveau als relatief nadelig worden beoordeeld, gezien de geringe voordelen bij gecontroleerde vergelijking tussen groepen.^{7,8} Aan de andere kant zou een individuele patiënt met keelpijn erop kunnen gokken dat hij wat eerder opknapt met antibiotica. De NHG-standaard Acute Keelpijn weegt deze factoren en adviseert de gulden middenweg te zoeken door pas bij een kans van meer dan 50 procent op een streptokokken-keelontsteking met antibiotica te behandelen.

NHG-standaard: compromis en consensus

Deze afwegingen zijn zoveel mogelijk gebaseerd op empirische gegevens, maar niet alle onderdelen van een NHG-standaard kunnen op deze manier worden onderbouwd, bij gebrek aan goed onderzoek. In dat geval wordt consensus nagestreefd. De vaststelling van NHG-standaarden kent daarom vele beoordelingen door een groot aantal huisartsen in de diverse fasen van de procedure.⁹ Dit zware accent op beoordeling, weging van feiten, compromis en consensus binnen de beroepsgroep maakt dat aan de NHG-standaarden bij voorbaat veel gewicht moet worden toegekend. Dit betekent niet dat toepassing van NHG-standaarden in alle situaties gepast is. Het betekent zelfs niet dat NHG-standaarden altijd 'goed' zijn: NHG-standaarden steken soms niet goed in elkaar en sommige huisartsen leveren optimale zorg door op onderdelen zeer afwijkend te handelen.

Gepast gebruik van NHG-standaarden: toetsgroepen

Een belangrijke indicatie voor toepassing van een NHG-standaard is een bijeenkomst van een intercollegiale toetsgroep, één van de belangrijkste ontwikkelingen in de Nederlandse huisartsgeneeskunde, ontstaan in 1978.¹⁰ Het gaat hier om individuele huisartsen met hun gewoonten, ervaringen, kennis en (on)zekerheden. Toetsgroepen kunnen gebruik maken van een bepaalde NHG-standaard, die als een agenda kan dienen voor het afspreken van een gefundeerd beleid.¹¹ Dit beleid zou idealiter prospectief

getoetst moeten worden met registratiegegevens uit de eigen praktijk, waarna de afspraken bestendigd of bijgesteld worden. In feite wordt in toetsgroepen de discussie die bij de vaststelling van een NHG-standaard is gevoerd nog eens dunnetjes overgedaan. Hierbij kan de interpretatie van het feitenmateriaal dat tot de richtlijn heeft geleid anders uitvallen in een toetsgroep dan in de NHG-standaard staat verwoord. Individuele normen en waarden van huisartsen, en van hun patiënten, zijn de elementen waaraan de richtlijn getoetst wordt door de groep huisartsen. Haalbaarheid en taaie gewoonten, van huisartsen en van patiënten, zijn vervolgens meer aardse argumenten om niet precies volgens de richtlijnen te handelen. Dit heeft te maken met de complexiteit van de huisartsgeneeskundige zorg.¹²

Richtlijnen over één onderwerp zijn door hun aard noodzakelijke simplificaties. Als zo'n monografie dan bovendien de differentiële diagnostiek afdoet met een paar zinnen, dan is de barrière voor toepassing in de praktijk wel erg hoog geworden. Daarom is enige bescheidenheid over de beïnvloedbaarheid van huisartsgeneeskundige zorg door NHG-standaarden zeker gepast.¹³

Afwijken van een NHG-standaard is dus geen kwestie van afwijken bij die ene unieke patiënt. Het gaat niet om de uitzondering op de regel. Afwijken van een NHG-standaard is normaal, omdat er geen norm-patiënt bestaat. Het gaat er om *met argumenten* af te wijken van een richtlijn, desnoods altijd. Zo bezien kan zelfs een onuitvoerbaar richtlijn een functie hebben.

Het doel van intercollegiale toetsing is een beargumenteerd en gedragen beleid af te spreken, en te evalueren. De NHG-standaard is niet een rode streep waar voorbij slecht handelen begint, en is géén gouden standaard. Het te allen tijde volledig volgen van NHG-standaarden door alle huisartsen is een onjuiste doelstelling.

De conclusie is dat NHG-standaarden voor toetsgroepen fungeren als een agenda voorzien van richtlijnen voor de discussie. Of de besluitenlijst van zo'n vergadering een betere kwaliteit van zorg oplevert, zal moeten worden aangetoond.

Waarom moeten toetsgroepen worden afgemeten?

De werkzaamheden van toetsgroepen dienen afgemeten te worden aan uitkomsten die voor de zorg belangrijk zijn. Bij keelpijn kan dat het aantal antibioticavoorschriften zijn, en bij diabetes het aantal goed ingestelde diabeten. Als blijkt dat er geen redenen of geen mogelijkheden zijn om de zorg te verbeteren, dan hoeft de nieuwste NHG-standaard natuurlijk niet op de agenda. Als een onderwerp wel op de agenda komt, dan is het niet eenvoudig vast te stellen wat belangrijke uitkomsten zijn en wat vervolgens verbeterde uitkomsten zijn. Zeker is dat het volgen van een

NHG-standaard op zich zelf géén goede zorg is, hoogstens een stap op weg. Het gaat immers niet om de richtlijn, maar om goed klinisch handelen.

Toetsingspakketten bevatten nogal eens de opdracht alle voorgeschreven criteria voor de diagnostiek te registreren op het spreekuur, van keelpijn bijvoorbeeld. Maar daar gaat het niet om. Het gaat erom of, in het geval van keelpijn, niet te veel patiënten ten onrechte antibiotica krijgen. Wat 'te veel' is, bepaalt de toetsgroep zelf, aan de hand van de NHG-standaard en eigen registratiegegevens.

Als we dit voorbeeld verder uitwerken, dan zou de uitkomst van de toetsing vastgesteld kunnen worden op basis van een registratie van alle keelpijn, alle gerichte diagnostiek en alle bijhorende farmacotherapie in het verloop van alle relevante ziektegeschiedenissen. Men kan dit meer of minder geavanceerd uitvoeren, of men kan een benadering van deze uitkomst zoeken. Een grove benadering bij keelpijn zou het percentage antibiotisch behandelde keelpijn kunnen zijn. Als men daar tevreden over is, en de collega's vergelijkbare uitkomsten hebben, dan is het onderwerp afgesloten. Wil men wel doorgaan, dan kan het zinvol zijn de diagnostiek te oefenen, te registreren en wederom de gekozen uitkomst te meten. Maar diagnostiek blijft een tussenstap, want het gaat bij het onderwerp keelpijn om adequaat antibioticagebruik, afgewogen tegen de belangen, verwachtingen en verlangens van de patiënt.

In dit kader is het niet zinvol huisartsen te vragen of ze denken dat ze altijd naar alle relevante klachten vragen. Het is evenmin zinvol te vragen of ze denken dat ze vaak antibiotica voorschrijven of niet. Resultaten van enquêtes zijn een kluwen van gerapporteerd gedrag en van de waarde die aan een standaard wordt toegekend. De gegevens die dergelijke opiniepeilingen opleveren, staan mijlenver af van de belangrijke uitkomsten van zorg.

Ongepast gebruik van NHG-standaarden

De toepassing van een NHG-standaard als rigide norm is ongepast. Op indicatie kan afgeweken worden van de richtlijn en dat gebeurt ook in de praktijk. *Boekhoorn* laat zien dat 75 procent van de antibiotisch behandelde patiënten met keelpijn niet voldoet aan de criteria, maar dat daar vaak plausibele redenen voor zijn.¹⁴ Of dit goede of niet goede redenen zijn, staat te bezien, maar het is niet verboden antibiotica voor te schrijven bij keelpijn. Een verbod op farmacotherapie bij minder dan vier criteria met de rechterhand op de NHG-standaard Keelpijn is ongepaste orthodoxie.¹⁵

Een ander voorbeeld is het gebruik van de NHG-standaard in regionale diabetesprojecten. De laatste tijd komen daarvoor nogal wat fondsen beschikbaar en dus worden project-coördinatoren, transmurale diabetes-ver-

pleegkundigen en de nodige stagiaires aangesteld. Deze medewerkers trekken 'het veld' in en vragen aan huisartsen en assistentes of ze de NHG-Standaard Diabetes Mellitus Type II kennen en wat ze zoal vinden van dit onderwerp. De doelstelling van deze enquête-activiteiten is beperkt, maar op zichzelf niet zinloos. Men wil namelijk iets te weten komen over de structuur en over het proces van de diabeteszorg. Zowel de structuur als de procesmatige kant zijn voorwaarden voor een goede uitkomst: goed ingestelde en gecontroleerde diabetespatiënten. De vragenlijsten moeten daar dan ook op gericht zijn.

Dit betekent dat een kennistoets een goede toets moet zijn.¹⁶ Een gestructureerd interview komt daaraan zelden tegemoet. Er hoort bovendien gecorrigeerd te worden voor sociaal wenselijke antwoorden, want niemand wil dom lijken. En natuurlijk mag men bij het afnemen van de kennistoets de diabetes-standaard opslaan, want dat wordt in de praktijk ook gedaan. Het resultaat kan leiden tot bijscholing, waarna weer een kennistoets kan volgen.

Een structuuranalyse van de praktijkvoering door middel van een gestructureerd interview en kaartenbakonderzoek naar een complete registratie van gegevens kan leiden tot het verbeteren van de infrastructuur. Die verbetering kan weer getoetst worden. Lacunes in de praktijkvoering op zichzelf, zoals *Konings* die laat zien, zijn geen vanzelfsprekende uitkomst van diabeteszorg. Want deze hele operatie betreffende diabeteskennis en praktijkvoering moet getoetst worden op zijn effect: meer goed ingestelde en gecontroleerde diabetespatiënten.

Deze werkzaamheden hebben niets meer met geneeskunde te maken. Artsen zijn voor zo'n operatie ongeschikt. Hiervoor is expertise nodig van onderwijskundigen, organisatiedeskundigen en professionele vragenlijstmakers. En als dit allemaal lege artis is uitgevoerd, dan nog is het een zijpaadje. Het gaat niet primair om competentie, attitudes, leerprocessen of meningen. Het gaat om de primaire uitkomst: de instelling van diabetes, afgemeten aan de hand van patiëntgebonden informatie als HbA1-waarden, compliantie of het welbevinden van patiënten. Uiteindelijk gaat het zelfs om uitkomsten als complicaties, morbiditeit en mortaliteit, want daar is het in de geneeskunde uiteindelijk om begonnen.

Beide voorbeelden betreffen ongepast gebruik van NHG-standaarden. In het eerste geval is de NHG-standaard opgevat als rigide standaard in plaats van als richtlijn. In het tweede geval is de NHG-standaard opgevat als iets wat je moet weten, in plaats van iets dat effect moet hebben op de zorg.

Kwaliteit van zorg meten met behulp van NHG-standaarden

NHG-standaarden hebben tot doel de zorg te verbeteren, als dat tenminste nodig is. Eigenlijk zou elke NHG-standaard vergezeld moeten gaan van relevante uitkomstmaten.

Voor de NHG-standaarden Keelpijn en Diabetes zijn hiervoor enkele suggesties gedaan. Idealiter zou een NHG-standaard vóór publicatie getoetst moeten worden op zo'n uitkomstmaat. Er is dan tenminste een uitkomst bekend als referentie voor toetsgroepen en als referentie voor wie de invoering van de NHG-standaard zelf als interventie in de gezondheidszorg beschouwt. Vaak zal een afgeleide uitkomst voldoende zijn, omdat de ideale uitkomst – morbiditeit en functioneren over lange tijd – onbereikbaar is.

De ontwikkelaars van een NHG-standaard verkeren in de ideale positie voorstellen te doen voor zulke afgeleide uitkomstmaten die nadrukkelijk praktisch uitvoerbaar en klinisch relevant moeten zijn. In het voorbeeld van de diabeteszorg kan men zich voorstellen dat niet de HbA1-waarden over een jaar, maar een gestegen prevalentie van goed gecontroleerde diabetespatiënten in een praktijk een zinvolle en praktische uitkomst is. Voor astma is misschien het aantal exacerbaties te verkiezen als uitkomstmaat, af te meten aan het aantal nachtelijke consulten en het aantal prednisonkuren.

Retrospectief dossieronderzoek en enquêtes onder huisartsen echter zijn zelden zinvolle uitkomstmaten, zoals *Konings* in zijn artikel in dit nummer laat zien. De gegevens hieruit staan veel te ver af van waar het om gaat, namelijk goede zorg.

Conclusie

NHG-standaarden zijn flexibele richtlijnen, waarvan bear-gumenteerde afgeweken kan worden. De beste indicatie voor NHG-standaarden is daarom het gebruik in toetsgroepen, omdat daar de afweging gemaakt wordt tussen richtlijn en praktijk. Een NHG-standaard is géén gouden standaard voor kwalitatief goed handelen. Het is daarom niet gepast huisartsen af te rekenen op het niet naar de letter uitvoeren van de richtlijnen. In meeste gevallen is het evenmin gepast structuur- en proceskenmerken van zorg te beschouwen als primaire uitkomstmaten van goede zorg. Als men desondanks in het kader van de evaluatie van zorg geïnteresseerd is in deze structuur- en proceskenmerken van zorg, dan is het niet eenvoudig deze kenmerken goed te evalueren en af te meten aan de primaire uitkomst. Die primaire uitkomst kan bestaan uit de morbiditeit, het welbevinden van patiënten, of afgeleiden daarvan.

Voorgesteld wordt NHG-standaarden te voorzien van praktische en zinvolle instrumenten voor de evaluatie van zorg ten dienste van toetsgroepen. Daar zijn NHG-standaarden namelijk voor bedoeld.

N.P. van Duijn
B. Meyboom-de Jong

- 1 De Melker RA, Kuyvenhoven MM, Van der Velden J. Voor-schrijf- en verwijsgedrag van huisartsen bij bovenste-lucht-weginfecties. *Huisarts Wet* 1993; 36: 7-10.
- 2 Konings GPJM, Wijkel D, Rutten GEHM. Lukt het werken volgens de NHG-Standaard diabetes mellitus type II? *Huisarts Wet* 1995; 38(1): 7-11.
- 3 Weijts MGG, Van der Horst FGEM. Algemeen verloskundige taken van de huisarts in het eerste trimester van de zwangerschap. *Huisarts Wet* 1994; 37: 566-8.
- 4 Eddy DM. Designing a practice policy; standards, guidelines, and options. *JAMA* 1990; 263: 3077, 3081, 3084.
- 5 Meyboom-de Jong B. Standaarden voor huisartsgeneeskundig handelen: een goed begin, maar geen einde. *Ned Tijdschr Geneesk* 1993; 137: 2124-6.
- 6 Tielens VCL. Standaarden: het gezicht van de huisarts. *Huisarts Wet* 1989; 32: 3.
- 7 Balder FA, Dagnelie CF, de Jong LJ, Kootte H. NHG-Standaard Acute Keelpijn. *Huisarts Wet* 1990; 33: 323-6.
- 8 Dagnelie CF. Sore throat in general practice [Dissertatie]. Utrecht: Universiteit Utrecht, 1994.
- 9 Thomas S. Standard setting in the Netherlands: impact of the human factor on guideline development. *Br J Gen Pract* 1994; 44: 242-3.
- 10 Van der Voort H. De invoering van toetsing. *Huisarts Wet* 1980; 23: 150-7.
- 11 De Maeseneer J, Meyboom-de Jong B. Niets is standaard in de huisartspraktijk (en zelfs dat niet) [Commentaar]. *Huisarts Wet* 1991; 34: 2-3, 5.
- 12 Lamberts H. Huisartsgeneeskundig handelen bij enkelklachten. *Huisarts Wet* 1991; 34: 35-9.
- 13 Lamberts H. Hoe beïnvloedbaar is de huisarts? *Huisarts Wet* 1981; 24: 371-8, 424-30.
- 14 Boekhoorn HCM, van Ree JW, Dubois V. Antibiotica bij keelpijn. *Huisarts Wet* 1992; 35: 70-1.
- 15 De Melker RA, Kuyvenhoven MM. Management of upper respiratory tract infection in Dutch general practice. *Brit J Gen Pract* 1991; 41: 504-7.
- 16 Pollemans M. Kennistoets bij huisartsen [Dissertatie]. Maastricht: Rijksuniversiteit Limburg, 1994.

