

Effecten van onderlinge toetsing

1. Spreekuurgedrag van huisartsen

R. GROL, H. MOKKINK EN F. SCHELLEVIS*

In hoeverre is deelname aan het programma voor onderlinge toetsing van het Nijmeegs Universitair Huisartsen Instituut van invloed op de werkwijze van huisartsen? Door middel van observaties en audio-opnamen van spreekuurcontacten is het handelen van 43 huisartsen zowel voor als na afloop van het toetsingsproject beoordeeld. De deelnemers bleken in een aantal opzichten meer volgens de geformuleerde criteria te zijn gaan werken, vooral bij de anamnese, het geven van voorlichting en adviezen, de globale oriëntatie en de verheldering van de hulpvraag aan het begin van het contact, de systematische opbouw van het consult en het actief stimuleren van de inbreng van de patiënt. Daarnaast was er bij de anamnese en het lichamenlijk onderzoek een kleine toename van het aantal 'overbodig' geachte handelingen. De veranderingen bleken het grootst bij de huisartsen die oorspronkelijk het laagst hadden gescoord op de verschillende criteria.

Inleiding

Zowel door de beroepsgroep als door de overheid wordt al jaren gepleit voor de invoering van onderlinge toetsing door huisartsen: men verwacht dat onderlinge toetsing kan bijdragen aan een verbetering van de kwaliteit van de huisartsgeneeskundige zorg. De vraag is echter of dit inderdaad het geval is. Onderzoek op dit terrein is in Nederland nog nauwelijks uitgevoerd en de resultaten van onderzoek in het buitenland zijn tegenstrijdig: onder bepaalde condities kunnen veranderingen optreden, maar de factoren die hierbij een rol spelen, zijn vaak niet duidelijk.¹⁻⁵

Sinds 1981 wordt door het Nijmeegs Universitair Huisartsen Instituut op grote schaal geëxperimenteerd met groepswijze onderlinge toetsing, gericht op een breed gebied van zowel het medisch als het niet-medisch handelen van de huisarts. Het feitelijke handelen in de spreekkamer wordt daarbij systematisch vergeleken met expliciete nor-

men of criteria voor adequate huisartsgeneeskundige zorg, die vooraf op basis van literatuurstudie en consensus-besprekingen zijn opgesteld.⁷⁻¹⁰ Doel van dit toetsingsproject is de deelnemers een aantal vaardigheden te leren op het gebied van onderlinge toetsing en een grotere bewustwording te bewerkstelligen ten aanzien van de eigen wijze van werken, met het streven daarin – voor zover nodig – correcties aan te brengen.

In een tweetal artikelen zullen we nagaan in hoeverre met name dit laatste is gelukt. In dit eerste artikel komt het medisch en niet-medisch handelen tijdens het spreekuur aan bod; het volgende artikel zal gewijd zijn aan de voorschrijf- en verwijsroutines van de deelnemers. Centraal staat de vraag, in hoeverre huisartsen door deelname aan toetsing meer in overeenstemming met een aantal normen voor adequate huisartsgeneeskundige zorg zijn gaan werken.

Methoden

Proefpersonen

Aan het project deden ongeveer 240 huisartsen mee. De hier beschreven deelstudie is uitgevoerd bij de deelnemers aan de eerste zes – van de in totaal 22 – toetsingsgroepen. Bij 43 van de 57 huisartsen uit deze groepen was het mogelijk gegevens te verzamelen over hun werkwijze in de spreekkamer 0,5-1 jaar vóór en 0,5-1 jaar na het project.

De deelnemers waren allen ervaren huisartsen die solistisch werkzaam zijn. Ze zijn geselecteerd op basis van het feit dat ze een groot aantal ziekenfondsverzekerden bij een tweetal regionale ziekenfondsen hebben; er is dus geen sprake van een aselechte steekproef. De huisartsen verschillen echter op een aantal variabelen (leeftijd, omvang en opbouw van de praktijk naar leeftijd en geslacht) niet veel van de landelijke populatie.

Praktische, financiële en methodologische overwegingen hebben ertoe geleid dat voor dit deel van het onderzoek geen controlegroep is gevormd.

Het toetsingsprogramma

Gekozen is voor onderlinge toetsing met vooraf opgestelde normen, en dus niet voor een toetsingsmethode waarin huisartsen informatie over de eigen werkwijze uitwisselen en de daarachter gelegen impliciete normen expliciet maken.¹¹⁻¹⁵ De deelnemers kwamen gedurende ongeveer anderhalf jaar maandelijks één hele dag bij elkaar in groepen van gemiddeld tien personen. Hen werd een intensief en gestructureerd toetsingsprogramma aangeboden. Door middel van een handleiding voor begeleiders werd de aanpak per groep zoveel mogelijk gestandaardiseerd,¹⁰ absolute standaardisering is natuurlijk, door de vele oncontroleerbare factoren, onmogelijk.

Zowel algemene huisartsgeneeskundige vaardigheden (gericht en systematisch werken, hanteren van de arts-patiënt relatie, somatisch en psychosociaal handelen) als het medisch-technisch handelen bij veel voorkomende klachten en aandoeningen werden onderling getoetst. De hierbij gehanteerde normen⁹ werden in de groepen ter discussie gesteld – en dus niet dwingend opgelegd – en gebruikt om eventuele inadequate werkwijzen op te sporen. In het begin was er een intensieve begeleiding, later werd toegewerkt naar volledige zelfstandigheid van de groep.

Beoordeling van het spreekkamergedrag

De beoordeling van de werkwijze vóór en na deelname aan het toetsingsproject vond plaats door middel van observaties in de spreekkamer en het opnemen en scoren van spreekuurcontacten op audioband; volgens Wakefield biedt deze methode de meeste informatie over aard en kwaliteit van het werk.¹⁶ Hierbij is gebruik gemaakt van een gestandaardiseerd beoordelingsinstrument, dat is ontwikkeld en getest binnen het onderzoeksproject Huisarts en Somatische Fixatie.¹⁷⁻²¹ Per huisarts werden in de voormeting 15-20 en in de nameting 13-15 consulten in de beoordeling betrokken. Het kader op bladzijde 100 geeft een overzicht van de variabelen waarop de werkwijze van de huisartsen is beoordeeld.

Het gebruikte instrument is zeer breed van opzet en meet een aantal aspecten van het huisartsgeneeskundig handelen, die in recente publikaties als belangrijk worden genoemd. Tussen de verschillende onderdelen van het instrument bestaan veel sterke en goed interpreteerbare samenhangen. Bovendien zijn er duidelijke samenhangen met voorschrijf- en verwijscijfers en vragen-

* Alle auteurs zijn verbonden aan het Nijmeegs Universitair Huisartsen Instituut.

lijstgegevens, zodat het instrument als redelijk valide beschouwd mag worden.

De normen in het meetinstrument zijn ontleend aan onze theorie over de preventie van somatische fixatie; zij zijn uiteindelijk gebaseerd op literatuurstudie en consensusbesprekingen.²² Uiteraard bieden deze normen geen absolute maatstaf voor goed huisartsgeneeskundig handelen: ze zijn maar ten dele onderbouwd met onderzoek, dat aantoonde dat de patiënt 'er beter van wordt' (*outcome*). Het wordt echter acceptabel geacht zich op het *proces* van de zorg te richten, zolang nog geen *outcome*-criteria beschikbaar zijn.^{23 24 13}

Voor de beoordeling van het medisch-technisch handelen is gebruik gemaakt van normen voor de aanpak van 24 veel voorkomende nieuwe klachten (hoofdpijn, keelpijn, hoesten, knieklachten, enz.) c.q. controles (hypertensie-, diabetes-, pilcontrole, enz.): zogenaamde medische protocollen. Daarin is aangegeven wat als adequaat handelen wordt gezien op het gebied van anamnese, onderzoek, voorlichting geven, terugbestellen, enz. Gemiddeld bevat elk protocol 30-40 'obligate' handelingen en daarnaast een aantal 'overbodige' handelingen.

Per huisarts zijn voor zowel de voor- als de nameting 400-600 handelingen in de beoordeling betrokken. Voor elke meting is berekend, hoeveel obligate handelingen waren verricht; dat aantal is vervolgens uitgedrukt als percentage van het *totaal* aantal obligate handelingen (dus het aantal obligate handelingen dat de huisarts tijdens de betreffende meting had kunnen verrichten). Tevens is per huisarts voor beide metingen het gemiddeld aantal overbodige handelingen per consult berekend.

Uit factoranalyse over de scores per huisarts over alle protocollen bleek dat er sprake was van één dimensie achter de scores. Hoog scoren op één protocol ging samen met hoog scoren op andere protocollen. Afgezien van een zekere variantie in het handelen op basis van verschil in klachten, bleken de protocollen dus tot op zekere hoogte uitwisselbaar. Een gemiddelde score over een groot aantal handelingen in een reeks consulten kan zo een indicatie geven van routines op medisch gebied.

De beoordeling vond plaats door vooraf getrainde en geïnstrueerde huisarts-beoordelaars. De interbetrouwbaarheid was – als men kijkt naar de uitkomsten van min of meer vergelijkbare beoordelingsstudies²⁵ – zeer goed te noemen: na correctie voor de kans op

overeenstemming bedroeg het overeenstemmingspercentage (kappa) .71 à .75.

Tevens is onderzocht of de aard van de voor elke meting gebruikte protocollen van invloed was op de mate van verandering. Daartoe is per huisarts nagegaan of bij de nameting meer gebruik is gemaakt van protocollen waarop men sneller een hoge score kan verwachten (als gevolg een kleiner totaal aantal obligate handelingen) dan in voormeting. Dit bleek niet het geval.

Voor de beoordeling van het niet-medisch handelen is gebruik gemaakt van gedetailleerde normen voor de vaardigheden 'gericht en systematisch werken', 'hanteren van de huisarts-patiënt relatie' en 'psychosociaal handelen'. Beoordeling van de voormeting vond plaats door zes ervaren gedragswetenschappers, van de nameting door drie getrainde huisartsen. Uiteraard werd voor beide metingen hetzelfde gestandaardiseerde instrument gebruikt. Het handelen wordt daarbij uitgedrukt in een aantal scores op een vijfpuntsschaal lopend van 'volledig in overeenstemming met de normen voor adequaat handelen' (1) tot 'totaal niet in overeenstemming met die normen' (5).

De interbetrouwbaarheid was wat minder hoog dan bij het medisch handelen (Kendall's tau bij de voormeting gemiddeld .30 à .50, bij de nameting .40 à .60). Dit ligt, gezien de complexiteit van de hulpverleningssituatie en de aard van de toegepaste criteria, voor de hand.¹⁶ Door de consulten van één huisarts over de beoordelaars te spreiden en door het uitvoeren van een controle- en correctieprocedure is ervoor gezorgd dat een eventuele vertekening in de resultaten zoveel mogelijk is voorkomen.

De onderzoeksvariabelen – The variables in the study

Medical performance

- Extent of adequate medical performance: extent to which a general practitioner performs medically necessary activities and omits careless and negligent performance.
- Extent of avoiding of superfluous medical performance: extent to which a general practitioner omits activities which in view of the present level of knowledge are not necessary.

Non-medical performance

- Aimed and systematic performance
 - initial phase: general orientation;
 - elucidation of question: explicitation of the patient's expectations;

Analyse

Om een indruk te krijgen van de veranderingen in het handelen van de deelnemers, zijn de gemiddelde scores over de beoordeelde consulten tijdens de voormeting (T1) en tijdens de nameting (T2) met elkaar vergeleken voor de gehele onderzoekspopulatie (n=43). Met behulp van een t-toets met gepaarde waarnemingen is bepaald welke waarde aan de gevonden verschillen kan worden toegekend. Het significantie-niveau is hierbij als maat gebruikt, ook al is er geen sprake van een aselechte steekproef en kan er niet zonder meer gegeneraliseerd worden. Verder is nagegaan hoeveel deelnemers (percentage) meer in overeenstemming met de normen zijn gaan werken.

Resultaten

Medisch handelen

Uit *tabel 1* blijkt dat het percentage verrichte obligate handelingen, over het totale consult gerekend, duidelijk is toegenomen: de meeste deelnemers zijn meer in overeenstemming met de normen gaan werken. Een dergelijke verschuiving vinden we zowel in de verhelderingsfase als in de beleidsfase. De grootste verandering is echter waar te nemen bij de anamnese en de medische voorlichting en advisering.

Deze verschuiving neemt echter niet weg, dat ook na deelname aan het toetsingsproject nog steeds gemiddeld 50 procent van de obligate handelingen *niet* wordt verricht. Ook zijn er duidelijke verschillen tussen de deelnemers: sommigen zijn niet, anderen zijn sterk veranderd. Vóór deelname aan het project scoorden 16 van de 43 huisartsen hoger dan 50 procent gemiddeld, na deelname waren dat er 24.

- systematic build-up of the encounter.
- Handling of the doctor-patient relation
 - listening, openness to the patient and taking him/her seriously;
 - active stimulation: actively involving the patient in decisions during the encounter;
 - explanation and information: giving the patient clear, understandable and direct information.
- Psychosocial performance
 - attention for psychosocial aspects of the complaints;
 - preventing that essential signals are disregarded.

Bij het overbodig medisch handelen gaat het over het algemeen om zeer kleine aantallen handelingen. Na het toetsingsproject zijn de deelnemers gemiddeld 0,4 overbodige handelingen per consult meer gaan verrichten (tabel 2). Nader analyse leert dat dit met name plaatsvindt in de verhelderingsfase en dan in het bijzonder een gevolg is van het stellen van meer anamnestiche vragen en het verrichten van meer handelingen tijdens het lichamelijk onderzoek.

Niet-medisch handelen

Ook wat het niet-medisch handelen betreft, zijn de deelnemers gemiddeld genomen meer in overeenstemming met de normen gaan werken (tabel 3). Het grootst zijn de verschillen bij de globale oriëntatie aan het begin van het consult, de hulpvraagverheldering en het actief stimuleren van de patiënt om mee te denken over wat er aan de hand is en wat er moet gebeuren. Ook voor de mate van systematiek in het consult, het geven van uitleg en informatie aan de

patiënt en het voorkomen dat er psychosociale signalen blijven liggen, is er een verandering in de richting van de gestelde criteria voor adequaat handelen.

Weinig verschil is er bij het luisteren naar en openstaan voor de patiënt (was al aan de positieve kant) en de aandacht voor psychosociale aspecten van de klachten. Op bepaalde punten zijn er duidelijke verschillen in de maten van verandering tussen de deelnemers onderling.

Tabel 1 – Table 1. Veranderingen in het medisch-technisch handelen (1): het percentage verrichte 'obligate' handelingen op T1 en T2, en het percentage deelnemers dat is veranderd in de beoogde richting – Changes in medical-technical performance (1): percentage of 'obligatory' activities performed at T1 and T2, and percentage of participants changed in the intended direction.

	T1	T2	Changed in the intended direction
Elucidation phase ^a	46.6	52.9	72%
– anamnesis ^a	50.4	58.1	70%
– physical examination	50.3	51.6	60%
Strategy phase ^a	40.0	46.4	67%
– information/advice ^a	37.7	45.9	67%
– asking to report back	44.6	48.6	63%
Entire encounter ^a	45.3	50.9	74%

^a Difference between T1 and T2 significant (t-test) $p \leq 0.01$.

Tabel 3 – Table 3. Veranderingen in het niet-medisch handelen: de gemiddelde scores^b op T1 en T2, en het percentage deelnemers dat is veranderd in de beoogde richting – Changes in non-medical performance: mean scores^b at T1 and T2 and percentage of participants changed in the intended direction.

	T1	T2	Changed in the intended direction
<i>Aimed and systematic performance</i>			
– general start ^a	3.8	3.1	91%
– elucidation of question ^a	4.4	3.3	100%
– systematic build-up ^a	2.4	2.0	81%
<i>Handling of doctor-patient relation</i>			
– listening and openness	2.5	2.4	42%
– active stimulation ^a	4.3	3.2	100%
– explaining and informing ^a	3.1	2.7	79%
<i>Psychosocial performance</i>			
– psychosocial attention	4.4	4.3	53%
– disregarding signals ^a	2.0	1.6	84%

^a Difference between T1 and T2 significant (t-test) $p \leq 0.01$.

^b Score 1: performance entirely up to standards; score 5: performance not at all up to standards.

Tabel 2 – Table 2. Veranderingen in het medisch-technisch handelen (2): het percentage verrichte 'overbodige' handelingen op T1 en T2 – Changes in medical-technical performance (2): percentage of 'superfluous' activities performed at T1 and T2.

	T1	T2
Elucidation phase ^a	0.6	0.9
Strategy phase	0.4	0.5
Entire encounter ^a	1.0	1.4

^a Difference between T1 and T2 significant (t-test) $p \leq 0.01$.

Tabel 4 – Table 4. Veranderingen (T1 → T2) bij deelnemers die op T1 respectievelijk laag en hoog scoorden – Changes (T1 → T2) in participants who at T1 scored low and high respectively.

	Low	High
<i>Medical performance</i>		
– adequate medical performance (elucidation phase) ^b	8.4%	2.3%
– adequate medical performance (strategy phase)	7.7%	5.1%
– avoidance of superfluous performance ^a	–	–0.7%
<i>Non-medical performance</i>		
– general start ^b	0.8	0.5
– elucidation of question	1.1	1.2
– systematic build-up ^a	0.7	0.2
– listening and openness	0.1	0.1
– active stimulation ^c	1.2	0.9
– explaining and informing ^a	0.7	0.1
– psychosocial attention	0.2	0.0
– disregarding signals ^a	0.7	0.1

^a Low-high difference significant (t-test) $p \leq 0.01$;

^b Low-high difference significant (t-test) $p \leq 0.05$;

^c Low-high difference significant (t-test) $p \leq 0.10$.

Verschillen in verandering tussen deelnemers

Uit de resultaten blijkt dat er – soms grote – verschillen tussen de deelnemers zijn ten aanzien van de doorgemaakte veranderingen. Van direct belang voor de interpretatie van de veranderingen is het uitgangsniveau van de deelnemers. De veranderingsscores van de huisartsen die vooraf hoog en de huisartsen die vooraf laag scoorden, zijn opgenomen in *tabel 4*. We zien dat zowel bij hoog- als bij laagscorders veranderingen in de beoogde richting zijn opgetreden, bij de laagscorders op bepaalde punten echter significant meer. Een uitzondering vormt het vermijden van overbodig medisch handelen. Huisartsen die vooraf weinig overbodig handelden, gaan meer overbodige handelingen verrichten dan huisartsen die al veel overbodig deden.

Beschouwing

De centrale vraag in dit onderzoek luidt: is de kwaliteit van de huisartsgeneeskundige zorg verbeterd door deelname aan onze groepsgewijze onderlinge toetsing?

Deze vraag is maar tot op zekere hoogte te beantwoorden. Er zijn nog zeer weinig harde criteria beschikbaar voor kwaliteit in de huisartsgeneeskundige zorg. Onderzoek en toetsing van het resultaat van de zorg op het welbevinden van patiënten brengt grote problemen met zich mee. Veel auteurs zijn daarom van mening dat de assumptie dat goed huisartsgeneeskundig handelen tot goede resultaten leidt, voorlopig een bruikbaar uitgangspunt is voor onderlinge toetsing.^{2 16 23 24} Ook al ontbreekt directe informatie over de gezondheid van de patiënt, dan is het toch zinnig en mogelijk het professionele handelen te toetsen.¹³

De door ons gehanteerde normen voor goed huisartsgeneeskundig handelen zijn zorgvuldig opgesteld vanuit de recente literatuur (wat niet betekent dat ze niet vatbaar zouden zijn voor discussie). Bij de bespreking in de groepen van huisartsen bleken ze ook grotendeels acceptabel. Meer in overeenstemming gaan werken met deze normen op het gebied van vraagverheldering, anamnese, geven van voorlichting aan patiënten, stimuleren van inbreng van patiënten, psychosociaal handelen terugbestellen, enz. zien we voorlopig dan ook als een ondersteuning van onze aanpak. Niettemin achten we verdere validering van de meetinstrumenten en de gehanteerde normen zeker noodza-

Synopsis

Grol R, Mokkink H, Schellevis F. Effects of peer group auditing.

1. Surgery hour behaviour of general practitioners. *Huisarts en Wetenschap* 1986; 29: 99-103, 124.

Introduction. Since 1981 the Nijmegen University Institute of General Practice has experimented on a large scale with peer group auditing. The aim of this auditing project is to teach participants some skills in peer auditing and to increase their awareness of their own performance. Two studies have been performed in order to establish the extent to which peer group auditing has caused participant general practitioners to begin working more in accordance with a number of standards of adequate general practice care.

Methods. The study described here was performed among the 43 participants in six auditing groups. From these general practitioners data were collected on their performance in surgery 6-12 months before and 6-12 months after the auditing project (T1 and T2 respectively).

Performance before and after participation in the auditing project was assessed by means of observations in the surgery and the recording and scoring of surgery hour encounters on audio-tape. Use was made of a standardized evaluation instrument developed and tested in the research project General Practitioner and Somatic Fixation. The variables on which the performance of the participants was tested, are shown on page 101.

Medical-technical performance was assessed with the aid of standards for the approach to 24 common new problems (headache, sore throat, cough, knee complaints, etc.) and follow-ups (hypertension, diabetes, oral contraception, etc.); so-called medical protocols. These protocols indicate what is regarded as adequate performance in terms of anamnesis, examination, information, asking to report back, etc. Each protocol comprises an average of 30-40 'obligatory' activities as well as a number of 'superfluous' activities. The number of obligatory activities was calculated for each general practitioner at T1 and T2; these numbers were then expressed as percentages of the total number of obligatory activities (i.e. the number of obligatory activities he could have performed at T1 and T2). Moreover, the median number of superfluous activities per encounter per general practitioner was calculated.

The evaluation was made by pre-trained and instructed general practitioners.

Non-medical performance was assessed with the aid of detailed standards for the

skills 'aimed and systematic performance', 'handling of doctor-patient relation' and 'psychosocial performance'. Measurements at T1 were performed by six experienced sociologists, and those at T2 by three trained general practitioners, of course using the same standardized instrument at T1 and T2. In this instrument performance is expressed in a number of scores on a five-point scale from 'entirely in accordance with the standards of adequate performance' (1) to 'totally at odds with these standards' (5).

To gain some impression of the changes in the performance of the participants, the median scores on the encounters assessed at T1 were compared with those at T2 for the entire study population (n=43). A t-test for paired observations was used to determine the value to be assigned to the differences found.

Results. *Table 1* shows that the percentage of obligatory activities performed over the entire encounter has clearly increased: most participants have begun to work more in accordance with the standards. Such a shift is observed both in the phase of elucidation and in the strategy phase. But the greatest change can be observed in anamnesis and medical information and advice. However, this shift cannot obscure the fact that an average of 50 percent of the obligatory activities still were not performed after the auditing project.

Superfluous medical performance generally concerned very small numbers of activities. At T2 the participants on average performed 0.4 superfluous activities more per encounter (*table 2*).

In terms of non-medical performance, too, the participants have on average begun to work more in accordance with the standards. Differences were most marked in general orientation at the start of the encounter, elucidation of the request for help, and active stimulation of the patient's participation in considering what is wrong and what should be done about it. Differences were small in listening and openness to the patient (was already quite positive) and in attention for the psychosocial aspects of complaints (*table 3*).

The results show that participants differed – sometimes markedly – in the changes they experienced. *Table 4* shows the changes in participants who respectively scored high and low at T1.

Keywords: Encounter; Family practice; Medical audit; Professional behaviour.

Address of authors: Nijmeegs Universitair Huisartsen Instituut, St. Annastraat 284, 6525 HC Nijmegen, The Netherlands.

kelijk. Enkele onderzoeken die hierop zijn gericht, zijn inmiddels van start gegaan of in voorbereiding.

De gevonden resultaten roepen verder een aantal vragen op.

In het toetsingsprogramma is veel aandacht besteed aan toetsing van de wijze van informatieverzameling, de gespreksvoering, het ingaan op vragen en denkbeelden van patiënten, het geven van uitleg en informatie en de systematiek in de hulpverlening. De deelnemers aanvaardden dit als essentiële aspecten van hun werk en de bereidheid om op deze punten te veranderen leek groot. Tot op zekere hoogte zien we dat gereflecteerd in de resultaten, zowel op medisch als op niet-medisch gebied. De deelnemers bevestigen ook in mondelinge en schriftelijke evaluaties dat ze op deze punten veranderingen hebben aangebracht in hun werkwijze.

De vraag is wat de veranderingen precies waard zijn. Zo zijn conclusies niet eenvoudig te trekken, omdat het bij dit soort onderzoek bijna onmogelijk is echt vergelijkbare controlegroepen samen te stellen.²⁶⁻²⁸

De veranderingen zijn op niet-medisch gebied het grootst. Men moet deze gegevens echter met enige voorzichtigheid bekijken, vanwege de mogelijke onbetrouwbaarheid in de beoordeling. Ondanks nauwkeurige formulering van de normen en uitvoerige training van de beoordelaars laat beoordeling van met name allerlei interpersoonlijke vaardigheden altijd een zekere mate van subjectieve interpretatie toe.¹⁶⁻²⁹

Op medisch gebied is sprake van een aantal 'significante' veranderingen. De verschillen tussen voor- en nameting zijn echter in absolute zin niet groot. Gemiddeld 50 procent van de obligate handelingen uit de protocollen worden ook ná toetsing nog steeds niet verricht.

Dit kan twee oorzaken hebben: de protocollen zijn te streng of er blijven grote lacunes bij sommige deelnemers bestaan. Waarschijnlijk spelen beide oorzaken een rol. De medische protocollen belichamen een ideaalbeeld, waaraan huisartsen hun eigen handelen kunnen spiegelen en waarmee zij lacunes kunnen opsporen.⁹ Ze geven een zo compleet mogelijk beeld van de te verrichten handelingen bij een klacht of aandoening. Voorgeschiedenis en voorkennis, het continuïteitsaspect in het arts-patiënt contact, zijn er echter niet in opgenomen. Dit feit zal ongetwijfeld verantwoordelijk zijn voor een deel van de afwijking van de protocollen.

Daartegenover staat een zeer grote

heterogeniteit in de scores, zowel vóór als ná de toetsing. Er zijn huisartsen die 60-70 procent en huisartsen die maar 30-40 procent van de obligate handelingen verrichten. Dit verschil kan niet uit de aard van de beoordeelde consulten worden verklaard; het wijst eerder op verschillen in medische routines.¹⁸ Uit ervaringen in de toetsingsgroepen en ook uit de literatuur blijkt dat vooral routines op medisch gebied zeer hardnekkig kunnen zijn.

Aparte aandacht verdient het overbodig medisch handelen. Uit nadere analyses blijkt dat meer volgens de normen van het niet-medisch handelen gaan werken samengaat met meer anamnestiche vragen en lichamelijk onderzoek, die niet per se nodig zijn bij een klacht. Wij denken dat de deelnemers aan toetsing gemiddeld genomen wat bewuster en minder 'minimalistisch' zijn gaan werken en dat dit ook meer overbodige diagnostische handelingen met zich heeft meegebracht. We vinden deze toename niet erg verontrustend, omdat het om zeer kleine aantallen gaat. Daarnaast lijkt het risico van somatische fixatie hierbij geringer dan bij het onnodig voorschrijven van medicijnen of onnodig verwijzen en terugbestellen.

Op de verschillen tussen huisartsen onderling is al gewezen. We zien dat met name huisartsen die vooraf laag scoorden op de normen, meer in overeenstemming met die normen zijn gaan werken. Een 'regressie naar het gemiddelde effect' kan hier een zekere rol gespeeld hebben. Dit is echter niet waarschijnlijk, aangezien ook de hoogscorders in beoogde richting zijn veranderd.

In voorbereiding is een analyse om na te gaan welke factoren hiervoor mogelijk verantwoordelijk zijn. Zo blijkt dat huisartsopleiders met betrekking tot de mate van voldoende medisch handelen veel minder zijn veranderd dan huisartsgasteren en perifere huisartsen. De laatste scoorden bij de voormeting gemiddeld onder het niveau van de opleiders, maar kwamen na toetsing op hetzelfde niveau. Bij het niet-medisch handelen daarentegen bleken de opleiders het meest veranderd. Op dit moment hebben we de indruk dat de meeste veranderingen daar optraden, waar ook de meeste verandering mogelijk was. Bovendien zit er ongetwijfeld een plafond aan wat men, gezien het tijdsbestek van het project en de mogelijkheden in de dagelijkse praktijk, aan veranderingen kan realiseren.

- ¹ Sluys E, Dopheide J, Van der Zee J. Overzichtsstudie onderzoek eerste lijn. Utrecht: Nivel, 1985.
- ² Fowkes F. Medical audit cycle: a review of methods and research in clinical practice. Med Educ 1982; 16: 228-38.
- ³ Sheldon M. Medical audit in general practice [Occasional paper 20]. London: Royal College of General Practitioners, 1982.
- ⁴ Haynes R, Davis D, McKibbon A, Tugwell P. A critical appraisal of the efficacy of continuing medical education. JAMA 1984; 251: 61-4.
- ⁵ Eisenberg J. Physician utilization. The state of research about physicians practice patterns. Med Care 1985; 23: 461-83.
- ⁶ Vervallen.
- ⁷ Grol R, Mesker P. Een programma voor onderlinge toetsing voor huisartsen. Huisarts en Wetenschap 1984; 27: 198-204.
- ⁸ Grol R, Mesker P. Bewustwording en verandering door onderlinge toetsing. Huisarts en Wetenschap 1984; 8: 270-4.
- ⁹ Grol R, Mesker P, red. Huisarts en toetsing. Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1984.
- ¹⁰ Grol R, Schellevis F. Een programma voor onderlinge toetsing: handleiding voor begeleiders. Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1985.
- ¹¹ Van de Rijdt T, Touw I, Vermue R. Werkboek deel A. Toetsen van somatisch handelen. Utrecht: O & O, 1985.
- ¹² Van Weel C, Van Zelst P. Het handelen van huisartsen bij luchtwegaandoeningen. Huisarts en Wetenschap 1982; 25(suppl 6): 35-9.
- ¹³ Lamberts H. Hoe beïnvloedbaar is de huisarts? Huisarts en Wetenschap 1981; 24: 371-8, 424-30.
- ¹⁴ Boeke A, Knottnerus J. Toetsing bij hoesten. Huisarts en Wetenschap 1982; 25(suppl 6): 40-3.
- ¹⁵ Watkins C. Care of patients with congestive heart failure. Update 1984; 1 sep: 355-64.
- ¹⁶ Wakefield J. Direct observation. In: Neufeld V, Norman G, eds. Assessing clinical competence. New York: Springer, 1985.
- ¹⁷ Mokkink H, Smits A, Grol R. Prevara: een observatie-instrument van somatische fixatie. Ned Tijdschr Psychol 1982; 37(1): 35-50.
- ¹⁸ Smits A, Mokkink H, Grol R, e.a. Spreekuurgedrag van huisartsen. Huisarts en Wetenschap 1985; 28: 121-5.
- ¹⁹ Mesker P, Mesker-Niesten J, Mokkink H. Pijn in de keel, een poging tot een protocollaire benadering. Huisarts en Wetenschap 1982; 25(suppl 6): 44-8.
- ²⁰ Smits A, Mokkink H, Mesker P. Preventie van somatische fixatie, deel 4: Prevara, een meetinstrument voor het handelen van de huisarts tijdens het spreekuur. Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1984.
- ²¹ Mesker-Niesten J, Mokkink H, Mesker P, e.a. Een protocollaire benadering van rugklachten. Huisarts en Wetenschap 1982; 25(suppl 6): 44-8.

Vervolg literatuur op pag. 124.

Vervolg van pag. 103

- ²² Grol R, red. Huisarts en somatische fixatie. Utrecht: Bohn Scheltema & Holkema, 1983.
- ²³ Shaw C. Aspects of audit. Br Med 1980; 280: 1256-8, 1314-6, 1361-3, 1443-6, 1509-11.
- ²⁴ Donabedian A. Promoting quality through evaluating the process of patient care. Med Care 1968; 6:181-202.
- ²⁵ Koran L. The reliability of clinical methods, data and judgements. New Engl J of Med 1975; 293: 642-6, 695-701.
- ²⁶ Albinsky M. Onderzoek en actie. Assen, Amsterdam: Van Gorcum, 1978.
- ²⁷ Smith N. The feasibility and desirability of experimental methods in evaluation. Evaluation and Program Planning 1980; 3: 251-6.
- ²⁸ Winickof R, Coltin R, Morgan R, et al. Improving physician performance through peer comparison feedback. Med Care 1984; 6: 527-34.
- ²⁹ Katz F, Snow R. Assessing health workers performance. A Manual for training and supervision [WHO Public Health Papers, nr. 72]. Geneva: World Health Organization, 1980.

Literatuur bij Van der Meer, pp. 104-6.

- ¹ Colenbrander M, Fehmen C, Hammers Th. Medische ethiek 1936. Amsterdam: Koninklijke Maatschappij tot Bevordering der Geneeskunst, 1936.
- ² Verbeek-Heida PM. Verslag van een onderzoek naar de bestaande onderlinge samenwerking van de huisartsen in de provincie Groningen. Groningen: Laboratorium voor Sociale Geneeskunde, 1967.
- ³ Gill K. Sociaal-verloskundige verkenningen [Dissertatie]. Nijmegen: Dekker & Van de Vegt, 1973.
- ⁴ Bruins CP, Noordhoff KH, Verdenius W. Onderlinge consultatie door partners in een groepspraktijk. Huisarts en Wetenschap 1973; 16: 372-4.
- ⁵ Lamberts H. Praten over patiënten: een kwantitatieve benadering van de communicatie binnen een multidisciplinair team. Huisarts en Wetenschap 1976; 19: 181-5.
- ⁶ Bremer GJ. Over het gebruik van de huisartsen anamneselijst bij vage, onduidelijke klachten. Huisarts en Wetenschap 1980; 23: 53-8.
- ⁷ Van der Werf GTh, Bremer GJ. Een klinische les? Ned Tijdschr Geneesk 1983; 127: 2265-8.
- ⁸ Ter Braak EM, Van der Werf GTh. Verwijzen of niet? Kijken naar interdokterverschillen als methode van toetsing. Huisarts en Wetenschap 1984; 27: 334-7.

Ledencongres 1985

Het dertigste Ledencongres van het Nederlands Huisartsen Genootschap werd gehouden op 15 november 1985, in het Congressentrum Ouwehand te Rhenen. Hieronder het verslag van deze vergadering, die door 90 leden werd bijgewoond.

Om 20.30 uur verwelkomt voorzitter *Bottema* de aanwezigen, in het bijzonder de heer Van der Voort, die sinds februari 1985 in functie is als directeur van het NHG. Omdat er nauwelijks gebruik is gemaakt van de mogelijkheid om de benodigde stukken voor dit Ledencongres op te vragen, worden de stukken ter plekke uitgereikt. Afgesproken wordt om de stukken voor het Ledencongres 1986 tijdig rond te sturen.

De voorzitter geeft een toelichting op de gebeurtenissen van het afgelopen jaar: zo was er het vertrek van twee Dagelijks Bestuursleden en van de oude directeur, de komst van de nieuwe directeur en de noodzaak om het financiële beheer van het NHG beter te regelen, teneinde een grotere beheersbaarheid te verkrijgen. Vervolgens wordt het verslag van de 29e Huishoudelijke vergadering onder dankzegging aan de notuliste gearresteerd.

Beleidsplan 1986

De voorzitter licht toe dat het NHG een rol zal gaan spelen in het opzetten van een toetsings- en nascholingscurriculum voor huisartsen.

De *Geus* verwijst naar het Beleidsplan bladzijde 8 onder deskundigheidsbevordering en kwaliteitsbevordering en vraagt of de SDH-opties gelijk zijn aan de opties van de PAOG en hoe de opstelling van het NHG hierin is. De voorzitter antwoordt, dat het NHG-bestuursstandpunt besproken is met het IOH. De SDH is sterk gericht op het perifere niveau. Er ontstaan verschillende verbindingen tussen SDH-structuur en PAOG via het meso-niveau. In de gezamenlijke werkgroep van SDH en NHG werd een goede afstemming tussen PAOG en SDH op meso-niveau belangrijk genoemd. De werkgroep werkt aan een rapport dat voor 1 januari 1986 zal uitkomen. De voorzitter vervolgt, dat het LHV-standpunt met betrekking tot de bemanning van het me-

so-niveau door het NHG als 'ontoereikend' wordt ervaren.

Vervolgens vraagt de voorzitter om commentaar op de andere beleidsvoor-nemens (beleidsplan DB 85-119). De aanwezigen leveren geen op- en aanmerkingen en de voorzitter stelt vast, dat hiermede het verslag gearresteerd is. Het bestuur zal de opdrachten uitwerken en daartoe de benodigde stappen ondernemen.

Financieel Jaarverslag 1984

Penningmeester *Weijers* licht toe, dat de toenemende activiteiten binnen het Genootschap een zeer groot beslag leggen op de financiële middelen. Het jaar 1984 werd met tekorten afgesloten.

Chavannes informeert of deze financiële tegenvaller veroorzaakt wordt door minder winst bij Huisartsenpers. De voorzitter antwoordt dat dit voor een groot deel het geval is, en dat de ontwikkeling rond Huisartsenpers daarnaast een zaak op zich is. Ter discussie staat het economische draagvlak van Huisartsenpers. Dit leidde onder andere tot de vraag of er een samenwerkingsverband gezocht dient te worden met andere uitgevers of dat er gezocht moet worden naar nieuwe investeringsmogelijkheden.

Raeven (CMP) verwijst naar de laatste pagina van de begroting 1986 en informeert naar het grote verschil in de CMP-uitgaven voor 1983 en 1984 (f 5.077,- en f 19.764,99). Ook vraagt spreker zich af waarom de NHG-commissies of werkgroepen in het lopende jaar geen financieel overzicht ontvangen. De penningmeester antwoordt dat in 1986 een drie maandelijks controle zal plaatsvinden en hij licht de redenen uitblijven van tussentijdse financiële rapportage naar commissies en werkgroepen toe. De NHG-administratie wordt momenteel verzorgd door een extern bureau, dat in 1986 voor tussentijdse rapportage zorgt. Voorts noemt de penningmeester het enorme verschil tussen CMP-uitgaven 1983 en 1984. Ter verduidelijking voor zowel CMP als de penningmeester NHG, wordt een gesprek toegezegd tussen DB en CMP.

Van der Voort benadrukt dat voor een goede budgetbeheersing nodig is, dat de declaraties voor de betreffende commissies en werkgroepen tijdig ingeleverd worden. Het NHG heeft nieuwe declaratieformulieren ontwikkeld, waarop