

Naar standaardisatie van het instrumentarium voor het meten van de gezondheidstoestand

MARIE-LOUISE ESSINK-BOT
GOUKE J. BONSEL

Essink-Bot ML, Bonsel GJ. Naar standaardisatie van het instrumentarium voor het meten van de gezondheidstoestand. Huisarts Wet 1995; 38(3): 117-21.

Samenvatting Gezondheid-gerelateerde kwaliteit van leven ('gezondheidstoestand') is, naast overlevingsduur en ziektespecifieke parameters, een belangrijke effectmaat voor de evaluatie van medisch handelen. Zes generieke meetinstrumenten – SIP, NHP, COOP/WONCA-kaarten, MOS-20/24, SF/RAND-36 en EuroQol – werden met elkaar vergeleken op grond van literatuur en gebruikersoordelen. De zes instrumenten bleken inhoudelijk zoveel te verschillen, dat 'vertalen' van resultaten van het ene naar het andere instrument niet mogelijk lijkt. Qua testeigenschappen hebben ze verder in specifieke onderzoekssituaties hun eigen voor- en nadelen. Enige standaardisatie in de meting van de gezondheidstoestand is echter gewenst. Veel kan worden gewonnen door één instrument in principe standaard te laten 'meelopen' in evaluatiestudies. De COOP/WONCA-kaarten worden aanbevolen als 'meeloper' voor de proefperiode 1994-1996.

M.L. Essink-Bot, arts-epidemioloog, Instituut Maatschappelijke Gezondheidszorg, Erasmus Universiteit Rotterdam, Postbus 1738, 3000 DR Rotterdam. Dr. G.J. Bonsel, sociaal-geneeskundige, Instituut Klinische Epidemiologie en Biostatistiek, Universiteit van Amsterdam.
Correspondentie: M.L. Essink-Bot.

Inleiding

In de gezondheidszorg bestaat een sterke behoefte aan inzicht in de effectiviteit van medische voorzieningen. De gebruikers willen weten van welke behandeling de beste resultaten kunnen worden verwacht, en welke instelling daarvoor het meest in aanmerking komt; aanbieders van zorg willen patiënten de best mogelijke behandeling bieden; beleidsmakers hebben behoefte aan informatie om beslissingen over de verdeling van schaarse middelen te ondersteunen.

Gezondheid-gerelateerde kwaliteit van leven ('gezondheidstoestand') is een van de uitkomstmaten waarmee kan worden nagegaan of patiënten als groep baat hebben bij interventies, en zo ja, hoeveel. In vervolg daarop kan een indicatie worden gegeven van wat een specifieke patiëntengroep te verwachten heeft van een specifieke interventie.

Voor meting van de gezondheidstoestand zijn in Nederland thans zes generieke meetinstrumenten beschikbaar:

- de Sickness Impact Profile (SIP);¹⁻³
- de Nottingham Health Profile (NHP);^{4,5}
- de COOP/WONCA-kaarten;^{6,7}
- het 20-item instrument uit de Medical Outcomes Study (MOS-20/24);^{8,9*}
- de MOS 36-item Short Form Health Survey (SF-36)/RAND 36-item Health survey 1.0 (RAND-36);^{10,11}
- het EuroQol-instrument.^{12,13}

De eerste vijf instrumenten worden elders in dit nummer besproken. Het EuroQol-instrument is recent ontwikkeld door een Europese werkgroep, 'to complement other quality-of-life measures and to facilitate the collection of a common dataset for reference purposes'.¹³ Er is één Nederlandse versie. Met dit instrument wordt gezondheidstoestand beschreven door middel van vijf items met steeds drie niveaus (geen problemen, enige problemen, veel problemen). Het onderscheidt zich van de andere instrumenten, doordat het de mogelijkheid biedt *beschrijvingen* van gezondheidstoestanden te koppelen aan

empirisch verzamelde *waarderingen* voor gezondheidstoestanden. Dergelijke waarderingen zijn van belang bij economische evaluaties (kosten-utiliteitsanalyses).^{12,13}

Om de onderzoeker enig houvast te bieden bij het maken van een keuze uit deze zes instrumenten, bespreken wij hier een tweetal aspecten:

- De verschillen en overeenkomsten tussen de instrumenten.
- De wenselijkheid en mogelijkheden van standaardisatie.

Verschillen en overeenkomsten

De in *tabel 1* gepresenteerde vergelijking is gebaseerd op inhoudelijke (dus niet statistische) overwegingen. Vragenlijstschaalen (respectievelijk vragenlijst-items voor vragenlijsten met één item per schaal) zijn op dezelfde regel geplaatst en als vergelijkbaar beschouwd als de items ongeveer naar hetzelfde domein verwijzen. Het psychische domein komt in alle instrumenten aan de orde. Dit geldt ook voor het rolfunctioneren, zij het wat minder in de NHP. Sociale isolatie [NHP] verwijst naar onze mening naar het psychische domein. Fysiek functioneren wordt verschillend geoperationaliseerd. Een schaal voor 'Pijn' is niet aanwezig in de SIP; in de COOP/WONCA-kaarten is een pijnkaart optioneel. Pijn/Andere klachten [EuroQol] is de enige vragenlijst die ook naar andere somatische sensaties dan pijn verwijst. Vitaliteit [SF/RAND-36 en MOS-24] is de enige schaal die verwijst naar 'positieve gezondheid' ('erg energiek'); de items van de Energie-schaal [NHP] zijn alle negatief geformuleerd ('altijd moe').

Naar aanleiding van deze tabel concluderen wij, dat overeenkomstige schaal-labels een verschillende inhoud kunnen vertegenwoordigen (dezelfde vlag, verschillende ladingen). Ook het omgekeerde (verschillende vlaggen, overeenkomstige ladingen) komt voor. Bij beschouwing van de 'fysiek' gelabelde schalen blijkt bijvoorbeeld dat Lichamelijke fitheid [COOP] en de schalen Fysiek functioneren in MOS-20/24 en SF/RAND-36 alle

* MOS-24 = MOS-20 + Vitaliteitsschaal [mondelinge mededeling Dr. J.C.J.M. de Haes].

drie primair over fysieke inspanningstolerantie gaan, dat Mobiliteit [EuroQol], Fysieke mobiliteit [NHP] en Lopen [SIP] alle drie het fysieke vermogen tot lopen als onderwerp hebben, en dat Mobiliteit/Verplaatsing [SIP] actieradius meet.

In tabel 2 worden de zes vragenlijsten

vergeleken op een aantal gebruikspecificaties en psychometrische eigenschappen. De opgegeven waarden voor 'afname-duur' gelden vermoedelijk voor relatief gezonde patiënten. Vragenlijsten blijken vaak getest te worden in relatief gemakkelijk toegankelijke patiëntengroepen (niet

al te ziek, niet al te oud, geen visusstoornissen en een tamelijk voorspelbaar chronisch ziektebeloop). Gegevens over het 'gedrag' van vragenlijsten bij ernstig zieken zijn zeldzaam.

Dat een vragenlijst is ontworpen voor een bepaalde diagnose- of leeftijdsgroep,

Tabel 1 Inhoudelijke vergelijking van dimensie-inhoud van SIP, NHP, COOP/WONCA-kaarten, MOS-20/24, SF/RAND-36 en EuroQol

SIP	NHP	COOP/WONCA	MOS-20	SF/RAND-36	EuroQol
-	-	Lichamelijke Fitheid	Lichamelijk Functioneren	Fysiek Functioneren	-
Lopen	Fysieke Mobiliteit	-	-	-	Mobiliteit
Mobiliteit en verplaatsing	-	-	-	-	-
Emotioneel gedrag	Emotionele Reacties + Sociale Isolatie	Gemoedstoestand	Geestelijke Gezondheid	Mentale Gezondheid	Angst/depressie
Huishouden en Activiteiten in en om het huis + Sociale Interacties + Werk + Recreatie en vrije tijd	-	Dagelijkse Bezigheden + Sociale Activiteiten	Rolfunctioneren + Sociaal Functioneren	Rolbeperkingen (fysiek probleem) + Rolbeperkingen (emotioneel probleem) + Sociaal Functioneren	Dagelijkse Activiteiten
-	Pijn	-	Pijn	Pijn	Pijn/Andere klachten
-	-	Algemene Gezondheid	Ervaren Gezondheid	Algemene Gezondheidsbeleving	Waardering eigen gezondheid
Lichaamsverzorging en beweging + Eten	-	-	*	†	Zelfzorg
-	Energie	-	Vitaliteit‡	Vitaliteit	-
Slapen/rusten	Slaap	-	-	-	-
-	-	Veranderingen in de gezondheidstoestand (2 weken)	-	Gezondheidsverandering (1 jaar)	-
Alertheid en intellectueel functioneren	-	-	-	-	-
Communicatie	-	-	-	-	-

* Items als onderdeel van 'Lichamelijk functioneren'. † Items als onderdeel van 'Fysiek functioneren'. ‡ Alleen in de MOS-24.

betekent niet op voorhand dat zij niet toepasbaar zou zijn bij andere diagnose- en leeftijdsgroepen. Belangrijke testeigenschappen (test-hertest-betrouwbaarheid, interne consistentie, validiteitsaspecten) zouden opnieuw bepaald moeten worden bij gebruik in een andere populatie. Kennis van de herkomst van een vragenlijst kan soms verhelderend werken; het feit dat de NHP oorspronkelijk ontworpen is voor het evalueren van heupvervanging bij artrose, verklaart bijvoorbeeld het grote aandeel van de items over 'lopen' in deze vragenlijst.¹⁴

Toepasbaarheid bij verschillende leeftijdsgroepen betekent niet alleen dat de vragen adequaat moeten zijn, maar ook dat de patiënten moeten beschikken over het cognitieve vermogen een vragenlijst in te vullen. Of de vragenlijst vervolgens valide resultaten oplevert voor een andere leeftijdsgroep, is een geheel andere vraag.

Het vergelijken van de vragenlijsten op aspecten van betrouwbaarheid (zoals test-hertest-betrouwbaarheid) en validiteit (met name begrips- en inhoudsvaliditeit), bleek moeilijk te realiseren. Testeigenschappen zijn soms zeer verschillend geoperationaliseerd. Gepubliceerde test-hertest-intervallen bleken bijvoorbeeld uiteen te lopen van een uur tot een jaar. Gegevens over betrouwbaarheid en validiteit van de

instrumenten konden zelden worden vergeleken, doordat de daarop gerichte onderzoeken waren uitgevoerd in zeer verschillende studiepopulaties.

Standaardisatie

Geen enkel onderzoek staat geheel op zichzelf. Onderzoek vindt plaats in een onderzoekstraditie, waaraan instrumenten en methoden worden ontleend. Het is de bedoeling dat onderzoeksresultaten gebruikt worden, bijvoorbeeld als informatie bij het nemen van beslissingen. Ten behoeve van de bruikbaarheid van gezondheidstoestandmeting moeten onderzoeksresultaten enigermate vergelijkbaar zijn.

Over het algemeen is de keuze van een onderzoeksdesign (meetinstrumenten, timing van de metingen, analysemethoden) afhankelijk van de specifieke onderzoeksvraagstelling. Vergelijkbaarheid kan alleen worden bereikt door de keuze van instrumenten en methoden in gezondheidstoestandmetingen op elkaar af te stemmen. Het is daarom gewenst de keuzevrijheid van individuele onderzoekers en onderzoeksgroepen enigszins in te perken en een minimale vorm van standaardisatie na te streven. Dit geldt zowel voor het meetinstrumentarium als voor aspecten van het onderzoeksdesign. In het

volgende beperken wij ons tot de standaardisatie van meetinstrumenten.

Ervan uitgaand dat de meeste vragenlijsten voor gezondheidstoestandmeting oorspronkelijk Engelstalig zijn, kunnen de volgende niveaus van standaardisatie worden onderscheiden:

- 1 Vaste regels om Nederlandse versies te maken van bestaande vragenlijsten; vrije keuze uit deze instrumenten.
- 2 Eén Nederlandse standaardversie van elk instrument; vrije keuze uit de instrumenten.
- 3 Eén Nederlandse standaardversie van elk instrument; beperking van de keuze door aanbeveling van één set instrumenten.
- 4 Eén Nederlandse standaardversie van elk instrument; aanbeveling om standaard in elk onderzoek één bepaalde vragenlijst op te nemen (de 'common core'- of 'meeloper'-gedachte).
- 5 Eén standaard vragenlijst; geen keuzemogelijkheden.

De opties 1 en 2 zouden al een belangrijke verbetering zijn ten opzichte van de bestaande situatie; zij zouden op zichzelf echter nog niet bijdragen tot inhoudelijke vergelijkbaarheid van onderzoeksresultaten tussen studies. Duidelijk is verder dat optie 5 ongewenst is.

Optie 4 lijkt haalbaar. Met een 'meelo-

Tabel 2 Vergelijking van eigenschappen van generieke vragenlijsten voor meting van gezondheidstoestand

	SIP	NHP	COOP/WONCA	MOS-20/24	SF/RAND-36	EuroQol
Geschikt voor schriftelijk invullen door patiënt zelf	ja	ja	ja	ja	ja	ja
Aantal items	136	38 + 7	6	20/24	36	6
Afnameduur (min.)	20	5	5	5	10	2
Aard items	ja/nee	ja/nee	5-puntschalen met plaatjes	>1 type (3 - 6-puntschalen)	>1 type (2 - 6-puntschalen)	3-puntsschalen + 'thermometer'
Beoogde leeftijdsgroep	volwassenen	volwassenen	≥14 jaar	volwassenen	≥16 jaar	≥12 jaar
Beoogde diagnose(n)	gevarieerd	artrose heup	diagnosen in huisartspraktijk	gevarieerd	gevarieerd	gevarieerd

per' wordt een minimale vorm van standaardisatie verkregen, die in elk geval de mogelijkheid biedt inhoudelijke vergelijkingen tussen patiëntengroepen te maken en de plaats van groepen op het spectrum van 'gezond' tot 'zeer ziek' te bepalen.

De mogelijkheden voor standaardisatie op het terrein van gezondheidstoestandmeting worden sinds 1992 onderzocht door de Werkgroep Onderzoek Gezondheidstoestandmeting (WOG – zie *kader*). De WOG wil op basis van consensus en de wetenschappelijke 'state of the art' aanbevelingen aan gebruikers doen om inhoudelijke vergelijking van onderzoeksresultaten mogelijk te maken. Het doel is *niet* het geven van pasklare recepten voor de meting van de gezondheidstoestand.

De WOG heeft een afspraak gemaakt over een 'meeloper'-instrument voor de periode van 1994-1996. Deze afspraak houdt in dat de leden van de WOG zullen proberen in deze periode het gekozen instrument op te nemen in elke evaluatiestudie waarbij zij betrokken zijn.

Bij de keuze van een 'meeloper' bleek, bij gegeven goede kwaliteit, de *lengte* een van de belangrijkste criteria. Omdat niet verwacht kan worden (en het ook niet wenselijk is) dat alle onderzoeksgroepen hun 'eigen' instrument zullen vergeten en massaal zullen overstappen op de gekozen 'meeloper', zal de 'meeloper' in het algemeen als extra instrument fungeren. De meeloper moet dus kort en eenvoudig zijn.

Daarmee vielen SIP, NHP en SF/RAND-36 af. Het EuroQol-instrument is weliswaar ontworpen vanuit de 'common core'-gedachte, maar er is nog te weinig over gepubliceerd. MOS-20/24 wordt gezien als een voorloper van de SF/RAND-36. Als 'meeloper' voor de komende drie jaar zijn derhalve de COOP/WONCA-kaarten gekozen. De plaatjes op deze kaarten zorgen voor afwisseling, waardoor de toevoeging van de kaarten niet hinderlijk is. Het nut van de COOP/WONCA-kaarten als meeloper zal in de proefperiode worden geëvalueerd. Bekeken zal worden:

- in hoeverre het lukt de COOP/WONCA-kaarten te laten meelopen in evaluatiestudies;

Werkgroep Onderzoek Gezondheidstoestandmeting (WOG)

Het initiatief tot oprichting van de WOG is genomen naar aanleiding van het door de Adviesgroep voor het Stimuleringsprogramma Gezondheidsonderzoek (SGO) gefinancierde onderzoeksprogramma 'Standaardisatie in MTA' (medische technology assessment). Het SGO-onderzoeksprogramma 'Standaardisatie in MTA' is uitgevoerd door de Rijksuniversiteit Groningen, de Rijksuniversiteit Limburg en de Erasmus Universiteit Rotterdam.

Leden van de WOG (per 1 januari 1995) zijn: Prof.dr. J. Passchier (*voorzitter*),^a M.L. Essink-Bot, arts-epidemioloog (*secretaris*),^b Dr. N.K. Aaronson,^c Drs. C.H. Bakker,^d Dr. G.J. Bonsel,^e Prof.dr. G.A.M. van den Bos,^f Dr. J. Braspenning,^g Drs. A.F. de Bruin,^h Dr. J.J. van Busschbach,ⁱ Dr. J.C.J.M. de Haes,^k Dr. R.A.M. Erdman,^a C. König-Zahn, arts-epidemioloog,^l Dr. G.I.J.M. Kempen,^m Prof.dr. P.J. van der Maas,^b Dr. M.P.M.H. Rutten-van Mólken,ⁿ Dr. R. Sanderman,^m Dr. M.A.G. Sprangers,^c Prof.dr. F.W.M.M. Touw-Otten,^o Dr. E.M. ten Vergert,^p Prof.dr. C. van Weel,^q Dr. W. Wijker^k en Dr. J. Wolf.^q

- a Instituut Medische Psychologie en Psychotherapie, Erasmus Universiteit Rotterdam.
- b Instituut Maatschappelijke Gezondheidszorg, Erasmus Universiteit Rotterdam.

- c Nederlands Kanker Instituut, Amsterdam.
- d Vakgroep Interne Geneeskunde / Reumatologie, Academisch Ziekenhuis Maastricht.
- e Instituut Klinische Epidemiologie en Biostatistiek, Universiteit van Amsterdam.
- f Instituut Sociale Geneeskunde, Universiteit van Amsterdam.
- g Vakgroep Medische Informatiekunde, Epidemiologie en Statistiek, Katholieke Universiteit Nijmegen.
- h Vakgroep Medische Sociologie, Rijksuniversiteit Limburg.
- i Instituut voor Medische Technology Assessment, Erasmus Universiteit Rotterdam.
- k Instituut voor Medische Psychologie, Universiteit van Amsterdam.
- l Vakgroep Sociale Geneeskunde, Huisartsgeneeskunde en Verpleeghuisgeneeskunde, Katholieke Universiteit Nijmegen.
- m Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen.
- n Vakgroep Economie van de Gezondheidszorg, Rijksuniversiteit Limburg.
- o Vakgroep Huisartsgeneeskunde, Universiteit Utrecht.
- p MTA-bureau, Academisch Ziekenhuis Groningen.
- q Nederlands Centrum Geestelijke Volksgezondheid, Utrecht.

- in hoeverre dit bijdraagt tot vergelijkbaarheid van studieresultaten;
- in hoeverre dit bijdraagt aan een grotere bruikbaarheid van studieresultaten voor beleidsdoeleinden.

Beschouwing

De Vragenlijst voor Onderzoek van de Ervaren Gezondheidstoestand (VOEG) maakte geen deel uit van dit overzicht. De VOEG wordt in Nederland in de CBS-Gezondheidsenquête gebruikt als onderdeel van de gezondheidmeting.

De VOEG – oorspronkelijk ontworpen

als indicator voor stress in industriële situaties¹⁵ – bestaat uit een lijst van 48 (lange versie) of 21 (korte versie) 'vage', voornamelijk lichamelijke klachten, bijvoorbeeld 'trillende handen', 'pijnlijke voeten', 'prikkend gevoel in de neus', 'benauwdheid op de borst'. Van Sonsbeek heeft aannemelijk gemaakt dat de VOEG bruikbaar is als indicator van somatische gezondheid.¹⁶ De VOEG vraagt naar klachten, niet naar de gevolgen daarvan voor het functioneren. Daarbij is de invalshoek voornamelijk somatisch. De VOEG is op die grond geen generiek meetinstrument voor gezondheidstoestand.

In dit overzicht is voornamelijk verwezen naar sleutelpublicaties over de Nederlandse versies van de besproken meetinstrumenten; via deze publicaties kan men gemakkelijk literatuur over de oorspronkelijke versies vinden. Daarnaast is gebruik gemaakt van de deskundigheid van de leden van de Werkgroep Onderzoek Gezondheidstoestandmeting (WOG). Vele van de hier besproken vragen zijn in de ontstaansgeschiedenis van de WOG uitvoerig aan de orde geweest.

Gezondheidstoestand als uitkomstmaat is in de zes besproken meetinstrumenten voor een deel overeenkomstig, voor een deel verschillend geoperationaliseerd. In met name de SIP en de COOP/WONCA-kaarten zijn het fysieke, het psychische en het sociale domein functioneel geoperationaliseerd – als de *gevolgen* van gezondheid(sproblemen) voor het dagelijks leven –, terwijl bijvoorbeeld de NHP meer elementen bevat van *ervaren* gezondheid. Inhoudelijk verschillen de zes meetinstrumenten zoveel, dat 'vertalen' van resultaten van het ene naar het andere instrument niet mogelijk is.

Door het ontbreken van vergelijkbare documentatie over de zes generieke vragenlijsten is het op dit moment niet mogelijk te zeggen of het ene instrument beter is dan het andere. Bovendien: wat is 'beter'? Of een lange vragenlijst die gevoelig is voor kleine veranderingen, beter of slechter is dan een korte en wellicht minder gevoelige vragenlijst, is afhankelijk van de omstandigheden van een onderzoek. Enige standaardisatie van methoden om de gezondheidstoestand te meten is echter gewenst ten behoeve van onderlinge vergelijkbaarheid.

Momenteel wordt een registratiepunt voor verricht en lopend onderzoek met de COOP/WONCA-kaarten opgezet (Dr. R. Sanderman, Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen, telefoon 050-636.251). Op basis van de hier aangemelde onderzoeken kan ook het gedrag van de COOP/WONCA-kaarten in diverse populaties nader worden gedocumenteerd. Ook de lezers van *Huisarts en Wetenschap* worden hierbij uitgenodigd mee te werken aan stan-

daardisatie met behulp van de meeloper en aan verder onderzoek naar eigenschappen van de COOP/WONCA-kaarten.

Literatuur

- 1 Luttik A, Jacobs HM, De Witte LP. De Sickness Impact Profile. Utrecht/Maastricht: Vakgroep Huisartsgeneeskunde / Instituut voor Revalidatievraagstukken, 1987.
- 2 De Melker RA, Touw-Otten F, Jacobs HM, Luttik A. De waarde van de 'sickness impact profile' als uitkomstmeting. *Ned Tijdschr Geneesk* 1990; 134: 946-8.
- 3 De Bruin AF, Diederiks JPM, De Witte LP, et al. The development of a short generic version of the Sickness Impact Profile. *J Clin Epidemiol* 1994; 47: 407-18.
- 4 Erdman RAM, Passchier J, Kooijman M, Stronks DL. The Dutch version of the Nottingham Health Profile. *Psychol Rep* 1993; 72: 1027-35.
- 5 Essink-Bot ML, Van Agt HME, Bonsel GJ. NHP of SIP: een vergelijkend onderzoek onder chronisch zieken. *Tijdschr Soc Gezondheidsz* 1992; 70: 152-9.
- 6 Van Weel C. Functional status in primary care: COOP/WONCA charts. *Disabil Rehabil* 1993; 15: 96-101.
- 7 Scholten JHG, Van Weel C. Functional status assessment in family practice. *Lelystad: Meditekst*, 1992.
- 8 Kempen GJIM. Het meten van de gezondheidstoestand van ouderen; een toepassing van de Nederlandse versie van de MOS-schaal. *Tijdschr Gerontol Geriatr* 1992; 23: 132-40.
- 9 Bindman AB, Keane D, Lurie N. Measuring health changes among severely ill patients. *Med Care* 1990; 28: 1142-51.
- 10 Van der Zee K, Sanderman R, Heyink J. De psychometrische kwaliteiten van de MOS 36-item Short Form Health Survey (SF-36) in een Nederlandse populatie. *Tijdschr Soc Gezondheidsz* 1993; 71: 183-91.
- 11 Van der Zee KI, Sanderman R. Het meten van de algemene gezondheidstoestand met de RAND-36 – een handleiding. Groningen: Noordelijk Centrum voor Gezondheidsvraagstukken, Rijksuniversiteit Groningen, 1994.
- 12 Essink-Bot ML, Stouthard MEA, Bonsel GJ. Generalizability of valuations on health states collected with the EuroQol questionnaire. *Health Economics* 1993; 2: 237-46.
- 13 EuroQol Group. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16: 199-208.
- 14 McDowell I, Martini CJM, Waugh W. A method for self-assessment of disability before and after hip replacement operations. *BMJ* 1978; 2: 231-46.
- 15 Dirken JM. Het meten van 'stress' in industriële situaties. Groningen: Wolters, 1967.
- 16 Van Sonsbeek JLA. De VOEG: klaaglijs of lijst met gezondheidsklachten? Den Haag: SDU/CBS, 1990. ■

Abstract

Essink-Bot M.L., Bonsel G.J. Towards standardisation of instruments for health status assessment. *Huisarts Wet* 1995; 38(3): 117-21.

Health-related quality of life, or health status, is an important outcome measure in medical evaluation research, apart from survival and disease-specific measures. Six generic instruments currently available and used in the Netherlands – Sickness Impact Profile, Nottingham Health Profile, COOP/WONCA charts, the 20-item instrument from the Medical Outcomes Study, the MOS 36-item Short Form Health Survey / RAND 36-item Health Survey 1.0 and the EuroQol instrument – are described and compared on the basis of literature data and experts' opinions. The contents of the six instruments appeared to be different to such an extent that direct 'translation' of results of one instrument to another is precluded. Each instrument has its own specific advantages and disadvantages if classical psychometric questionnaire properties and test behaviour are regarded. Comparable information on testing properties appeared to be scarcely available. It is argued that a minimal level of standardisation of health status assessment is desirable from the viewpoint of comparability, and thus usability of study results. Opportunities for standardisation in the Netherlands are being explored in the Dutch Working Group on Health Status Assessment. A standard 'common core' instrument to complement other measures in any health status assessment provides at least some comparability. The Dutch Working Group recommends the use of the COOP/WONCA charts as a 'common core' instrument for a restricted period (1994-1996).

Key words Generic instruments; Health status assessment; Standardization.

Correspondence M.L. Essink-Bot MD, Department of Public Health, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands.