

De Nederlandse versie van de Sickness Impact Profile in huisartsgeneeskundig onderzoek

H.M. JACOBS
F.W.M.M. TOUW-OTTEN

Jacobs HM, Touw-Otten FWMM. De Nederlandse versie van de Sickness Impact Profile in huisartsgeneeskundig onderzoek. *Huisarts Wet* 1995; 38(3): 122-8.

Samenvatting Met generieke vragenlijsten zoals de Sickness Impact Profile (SIP) kunnen effecten van medische zorg uitgedrukt worden in termen van veranderingen in het dagelijks functioneren van patiënten. De SIP evalueert in 136 uitspraken de invloed van ziekte op het dagelijks functioneren over twaalf verschillende dimensies van de gezondheidstoestand. De SIP kan worden afgenomen door een interviewer in een vraagsprek, of door de patiënt zelf worden ingevuld, na een mondelinge of schriftelijke instructie. De Nederlandse versie is ontwikkeld in 1985-1987 en vertoont qua psychometrische eigenschappen een sterke overeenkomst met de originele Amerikaanse SIP. Sindsdien is en wordt het instrument veelvuldig in onderzoek gebruikt, zowel binnen als buiten de huisartsgeneeskunde. De patiënt-acceptatie is goed, maar bij oudere patiënten en bij vervolgregistraties kan de afnameduur van 25 tot 30 minuten een probleem zijn. De SIP onderscheidt goed tussen het functioneren van patiënten met ernstige en minder ernstige aandoeningen. Voor longitudinale toepassingen waarbij ziekte-specifieke in plaats van algemene functioneringsaspecten relevant zijn, is de SIP geen goede keus.

Vakgroep Huisartsgeneeskunde, Universiteit Utrecht, Universiteitsweg 100, 3584 CG Utrecht.

Dr. H.M. Jacobs, socioloog;
Prof.dr. F.W.M.M. Touw-Otten, socioloog,
hoogleraar huisartsgeneeskunde.
Correspondentie: Dr. H.M. Jacobs.

Inleiding

Verbetering van de functionele gezondheidstoestand is een belangrijke doelstelling van de huisartsgeneeskunde. Daartoe is een breed scala aan meetinstrumenten ontwikkeld: vragenlijsten waarmee de patiënt informatie vastlegt over één of meer dimensies van zijn toestand.^{1,2} Aan de hand van deze *klinimetriche indices* kunnen artsen, patiënten en onderzoekers de effecten van de medische zorg uitdrukken in termen van veranderingen in de gezondheidstoestand.^{3,4}

Eén van deze meetinstrumenten is de Sickness Impact Profile (SIP).⁵ De SIP is een generieke, multifunctionele vragenlijst om de invloed van ziekte op dagelijkse bezigheden en activiteiten vast te stellen (*tabel 1*).⁶ De patiënt wordt gevraagd naar feitelijk gedrag en niet naar symptomen, ziekteverschijnselen of ziektebeleving. Het instrument bestaat uit 136 uitspraken, verdeeld in twaalf categorieën die elk een specifiek gebied van dagelijkse bezigheden en activiteiten omvatten. Alle uitspraken zijn voorzien van een gewicht, zodat de ernst van gedragsverstoringen gedifferentieerd wordt weergegeven.

De scores kunnen zowel per categorie als voor de dimensies fysiek en psychosociaal functioneren worden vastgesteld; bovendien kan een totaalscore voor het algemeen functioneren worden bepaald. De scores worden berekend door de wegingsfactoren van de bevestigde uitspraken bij elkaar op te tellen en vervolgens te percenteren; hoe hoger de score, des te sterker de verstoring van het functioneren.

De meeteigenschappen van het originele instrument zijn uitvoerig onderzocht en beschreven;⁷⁻¹² in *tabel 2* zijn de belangrijkste conclusies samengevat.

In dit artikel bespreken wij de validiteit van de Nederlandse versie van de SIP en de ervaringen met dit instrument bij toepassingen in klinimetric, huisartsgeneeskundig onderzoek.

Nederlandse versie

De Nederlandse versie van de SIP is ontwikkeld in 1985-1987. Eerst werd nage-

gaan in hoeverre deze versie qua semantiek en inhoud gelijkwaardig aan het oorspronkelijke instrument was. Vervolgens werd getest in hoeverre de nieuwe versie technisch overeenstemde, bijvoorbeeld ten aanzien van de differentiatie naar ernstgraad tussen de afzonderlijke uitspraken. Een derde vergelijkingscriterium had betrekking op de overeenstemming in meetresultaten onder vergelijkbare condities. Tenslotte werd een uitspraak gedaan over de mate waarin de Nederlandse variant ook qua theoretisch concept overeenkwam met het oorspronkelijke instrument.^{13,14} Deze Nederlandse versie is en wordt veelvuldig toegepast in klinimetric onderzoek in de huisartsgeneeskunde,¹⁵⁻²² in de kliniek²³⁻²⁸ en in de revalidatiegeneeskunde.²⁹⁻³²

Analyse van bevindingen uit een deel van deze onderzoeken bevestigen dat de Nederlandse SIP qua psychometrische eigenschappen een sterke overeenkomst vertoont met de originele, Amerikaanse SIP. Vergelijkbare betrouwbaarheidscoëfficiënten werden gevonden voor de interne consistentie van de schalen²⁹ en subschalen, uitgezonderd een lagere interne consistentie in de Nederlandse categorie 'Slapen, Rusten'.²⁶ Ook voor de reproduceerbaarheid van de scores komen de Nederlandse uitkomsten sterk overeen met die van de originele SIP.^{32,33}

Bij een grootschalige toepassing van de SIP onder Nederlandse intensive-care-patiënten werd bij factoranalyse de door *Bergner et al.* beschreven theoretische factorstructuur van de SIP gevonden, terwijl in eigen onderzoek bij patiënten met niet-specifieke buikklachten de validiteit van de psychosociale dimensie bevestigd werd.

Patiënt-acceptatie en toepassing

De meeste patiënten reageren positief op de SIP. Slechts een enkeling geeft aan dat de voorgelegde beperkingen slecht aansluiten bij de eigen gezondheidstoestand;³⁴ daarbij gaat het met name om patiënten die geen van de 136 beschreven verstoringen op zichzelf van toepassing

Tabel 1 Inhoud en structuur van de Sickness Impact Profile

Dimensie/categorie (afkorting)	Aantal uitspraken	Bijdrage totaalscore	Dimensie/categorie (afkorting)	Aantal uitspraken	Bijdrage totaalscore
<i>Fysiek functioneren</i>			<i>Overig</i>		
- lopen (a)	12	8,4%	- werk (w)	9	5,1%
- lichaamsverzorging en beweging (bcm)	23	20,0%	- slapen en rusten (sr)	7	5,0%
- mobiliteit/verplaatsing (m)	10	7,2%	- huishouden en andere bezigheden in en om het huis (hm)	10	6,7%
<i>Psychosociaal functioneren</i>			- recreatie en vrije tijd (rp)	8	4,2%
- emotioneel gedrag (eb)	9	7,0%	- eten (e)	9	7,0%
- sociale interactie (si)	20	14,5%			
- alertheid/intellectueel functioneren (ab)	10	7,7%	Totale SIP-lijst	136	100,0%
- communicatie (c)	9	7,2%			

Tabel 2 De belangrijkste psychometrische eigenschappen van de Sickness Impact Profile**Validiteit**

- *Bewijs van de construct validiteit*
Een uitgebreide op gedragsevaluaties gebaseerd meetinstrument met duidelijk onderscheidende psychosociale en fysieke dimensies.
- *Inhouds validiteit van de gezondheidsdimensies*
Wat betreft de psychosociale dimensie bestaan er slechts middelmatige correlaties met traditionele depressie en emotie schalen. Hierdoor kan het zijn dat de SIP onvoldoende gevoeligheid bezit als dergelijke subjectieve gemoedstoestanden van belang zijn bij de evaluatie van patiënten uitkomsten.
- *Multitrait- Multimethod benadering*
Voldoende convergente en divergente validiteit.
- *Known-group validiteit*
Het instrument slaagt er goed in bij een breed scala aan ziekte categorieën verschillende niveau's te onderscheiden in psychosociaal en fysiek dysfunctioneren.

Cross-Culturele versies

Er bestaan verschillende internationale versies.

Antwoordcategorieën

Ja/ Neen antwoorden

Schalingskenmerken

- *Dimensionaliteit*
Twaalf verschillende dimensies.
- *Items per dimensie*
Multi item schalen, in totaal 136 items.

Index versus profiel

Voor het instrument kunnen afzonderlijke indexscores worden berekend, terwijl tevens een totale profiel score berekend kan worden.

Registratie/uitvoerbaarheid

Bij afname door een interviewer is een goede interviewer instructie noodzakelijk. Het afnemen /invullen van de SIP duurt tussen de 20 en 30 minuten. Acceptatie door patiënten is goed. Speciaal bij vervolgaafnames moet aandacht gegeven worden aan het voorkomen van response bias. De lengte beperkt de toepassingsmogelijkheden, echter door een goede organisatie van de dataverzameling kan dit worden ondervangen.

Betrouwbaarheid

- *Reproduceerbaarheid, test/hertest*
Een hoge reproduceerbaarheid voor de totale, psychosociale en fysieke index scores, lage reproduceerbaarheid voor de individuele itemscores. Een middelmatige reproduceerbaarheid in geval van administratie door de patiënt zelf.
- *Interne consistentie*
De interne consistentie van de 12 dimensies is middelmatig tot hoog met uitzondering van de de categorie 'Eten'.
- *Responsiviteit*
De gevoeligheid van de SIP om verschuivingen in de algemene functionele gezondheidstoestand ten gevolge van veranderingen in het klinisch beloop betrouwbaar vast te leggen is onvoldoende uitgekristaliseerd.
Het instrument lijkt onvoldoende responsief als uitkomst parameter voor ziektespecifieke functionele verstoringen.

achten. Verder blijken patiënten zonder functionele beperkingen bij vervolgmetingen minder geneigd de SIP volledig in te vullen dan patiënten die nog klachten hebben. Een ander probleem is het feit, dat oudere patiënten bepaalde beperkingen niet altijd ervaren als gevolg van een verslechterde fysieke gezondheidstoestand, maar als normale verschijnselen die bij het verouderingsproces behoren. Hierdoor worden allerlei verstoringen in het dagelijks functioneren niet altijd als zodanig bevestigd. Rothman *et al.* stellen daarom voor aan de instructie een omschrijving van gezond-zijn toe te voegen, waarin aan dit soort problemen aandacht wordt besteed.¹⁰

De duur van de registratie – 25 tot 30 minuten – is meestal geen probleem, behalve bij sommige oudere patiënten, met name als de SIP wordt afgenomen in combinatie met andere klinimetrische lijsten.²²

De SIP kan worden afgenomen door een interviewer in een vraaggesprek, of door de patiënt zelf worden ingevuld, na een mondelinge of schriftelijke instructie.³⁶

Hulsebos *et al.* vonden duidelijke verschillen in het algemeen functioneren bij drie vergelijkbare patiëntengroepen waarbij de SIP op verschillende wijzen was afgenomen. Hoewel de mogelijkheid van klinische verschillen niet geheel kon worden uitgesloten, was duidelijk dat de wijze waarop de SIP was afgenomen, van invloed was op de uitkomsten.³⁵ In longitudinaal onderzoek met verschillende vervolgmetingen moet de onderzoeker hiermee rekening houden (tabel 3).

Als de patiënt ernstig ziek is, zal hij niet altijd zelf de vragen kunnen beantwoorden, of is het ethisch niet verantwoord de metingen uit te voeren (bijvoorbeeld in geval van een ziekenhuisopname). Met name in longitudinaal onderzoek kan dit leiden tot zo'n sterke patiëntenuitval, dat daardoor de generaliseerbaarheid van de uitkomsten wordt aangetast. Alternatieven zijn dan het werken met vervangende scores of met 'proxies',^{37,38} bijvoorbeeld door de partner of de behandelend arts de vragen te stellen. In het onderzoek van Hulsebos *et al.* bleek echter, dat dit niet in alle gevallen tot antwoorden leidt. Bovendien

kunnen bij het gebruik van proxies systematische vertekeningen optreden.^{39,40}

Variatie in de functionele toestand

De SIP onderscheidt goed tussen ernstige en minder ernstige aandoeningen. Uit tabel 4 blijkt dat bij patiënten met alledaagse ziekten (niet-specifieke longziekten, buikproblematiek) lichte verstoringen in de functionele toestand geconstateerd werden. Daarnaast verschilt de functionele toestand van deze patiënten van de functionele toestand van gezonde personen en van open populaties. Verder blijkt dat bij patiënten met ernstig invaliderende aandoeningen, zoals CVA of reumatische artritis, overeenkomstig zwaardere verstoringen in het dagelijks functioneren worden vastgesteld. Patiënten met gewone ziekten – Cara, diabetes mellitus, vage buikklachten – ondervinden meestal slechts geringe verstoringen in hun algemeen functioneren; hierdoor worden bij deze groepen scores gevonden die over het algemeen dicht bij elkaar liggen.¹⁵

Tabel 3 Richtlijnen voor het gebruik van de SIP in huisartsgeneeskundig onderzoek

Wijze van afname

- De keuze voor schriftelijke afname door de patiënt zelf betekent het accepteren van een hogere random error in de scores. Mondelinge instructie door een praktijkassistente of per telefoon zal dit foutenpercentage verlagen, terwijl bij vervolgmetingen het confronteren met de baselijn scores hetzelfde positieve effect kan opleveren.
- De hoogte van de random error in SIP scores verkregen door mondelinge afname zal alleen laag zijn als de interviewers goed geïnstrueerd zijn en het veldwerk op een betrouwbare wijze is uitgevoerd.
- Het door elkaar gebruiken, switchen van verschillende registratie methoden in een studie moet alleen gebeuren na een expliciete afweging van de mogelijke beïnvloeding op de gezondheidsuitkomsten van de patiënten.

Uitspraken

- Bij patiënten zonder of met alleen minimale functionele verstoringen verbetert de acceptatie van de SIP door het niet oplezen van bepaalde uitspraken die evident niet van toepassing zijn, of het toevoegen van een schriftelijke toelichting bij uitspraken.

- Het nadrukkelijk herhalen van de instructie, dat veranderingen gerelateerd dienen te zijn aan de gezondheid van de patiënt vermindert het foutenpercentage. Dit geldt met name voor de categorieën 'Sociale Interactie' en 'Emotioneel functioneren'.

Organisatie van het veldwerk

- In een studie waarbij uitgebreide dataverzameling plaats vindt is het ondoenlijk om de SIP in de praktijk af te nemen. Het is gebleken dat in dergelijke situaties praktijkassistentes betrouwbare en toegewijde assistenten zijn wat betreft de coördinatie van de veldwerkzaamheden.
- Het structureel monitoren van het verloop van de dataverzameling bevordert het bereiken van een maximale response.
- Korte en directe communicatie lijnen tussen onderzoeksgroep en zowel patiënten, artsen en praktijkassistentes zorgt voor een blijvend hoge betrokkenheid, die de dataverzameling ten goede komt (lage drop-out).

Dit laatste heeft duidelijke consequenties. Indien de meerderheid van de patiënten al bij de aanvang van een studie een niet-verstoord algemeen functioneren laat zien, is evaluatie van interventies bij voorbaat zinloos: een verbetering van de scores bij vervolgmetingen is in die gevallen immers onmogelijk.⁴¹ Een tweede consequentie heeft betrekking op de statistische analyse van de uitkomsten. Veel statistische technieken zijn gebaseerd op de veronderstelling dat de variabelen min of meer normaal verdeeld zijn. Dit betekent dat, uitgaande van het gemiddelde gezondheidsniveau in de onderzoeksgroep, de overige gezondheidsscores symmetrisch over de schaal zijn verspreid, hetgeen hier niet het geval

is. Bij de analyses dienen dan ook veelal verdelingsvrije statistische methoden te worden gebruikt, of moeten de data getransformeerd worden naar een meer symmetrische scoreverdeling. Dit probleem doet zich overigens ook voor bij het gebruik van andere algemene indices.

Veranderingen in functionele toestand en klinisch beloop

Bij onderzoek naar de effectiviteit van bepaalde interventies staat de vraag centraal in hoeverre veranderingen door de tijd betrouwbaar en overeenkomstig het klinische beloop worden geregistreerd met de gebruikte klinimetrische indices. Dit

aspect – de responsiviteit van een klinimetrisch instrument – is voor de SIP slechts weinig onderzocht.

De responsiviteit kan worden bepaald door de waargenomen veranderingen in de functionele-toestandsscores te onderscheiden naar toevalsfouten en werkelijke verschuivingen. Deze kunnen vervolgens afgezet worden tegen het gedeelte in de functionele toestand dat werkelijk stabiel is gebleven. Met behulp van variantie- en covariantie-analyse kunnen vervolgens twee verschillende betrouwbaarheidsmaten worden berekend: een intra-klassereproduceerbaarheidscoëfficiënt en een responsiviteitscoëfficiënt (tabel 5).

In eigen onderzoek naar het handelen

Tabel 4 Gegevens over de functionele gezondheidstoestand in verschillende populaties

Toepassingsgebied/ SIP schalen	Gemiddelde fysieke-dimensie-score	Gemiddelde psychosociale-dimensie-score	Gemiddelde totaalscore	Omvang populatie
Gezonde controles, 18-75 jr ⁴⁵	0,3	0,9	0,8	239
Zf-verzekerden in 5 huisartspraktijken, 18-75 jr ⁴⁵	2,3	3,5	3,4	594
Open populatie uit één huisartspraktijk, 64-95 jr ²³	7,0	5,1	6,8	232
Niet specifieke buikklachten, leeftijd 18-75 jr ¹⁵	2,5	7,2	5,9	980
Chronische niet specifieke longziekten, leeftijd 40-84 jr	5,6	6,4	6,8	70
Chronische hemodialyse patiënten ²⁶	8,9	13,0	12,3	34
Reumatische artritis ³⁰	16,3	9,5	14,2	372
CVA populatie 8 weken na het accident ²³	21,8	13,4	20,1	111

Tabel 5 Enkele statistieken

Intra-klasse-responsiviteitscoëfficiënt.

Deze maat geeft de betrouwbaarheid aan waarmee een instrument behandelingseffecten tussen groepen patiënten vastlegt. Veranderingen in de functionele gezondheidstoestand tussen groepen patiënten worden beschreven naar de proportie van de variantie in de functionele gezondheidstoestandsscores dat statistisch verklaard kan worden door de verschillen in klinische uitkomsten. Bijvoorbeeld het meten van verschillende klinische effecten tussen diagnostische categorieën of tussen behandelde en niet behandelde patiënten.

Intra-klasse-reproduceerbaarheidscoëfficiënt.

Deze coëfficiënt geeft weer in welke mate aanwezige verschillen in de functionele gezondheidstoestand tussen patiënten door de tijd heen stabiel aanwezig blijven.

Effect-grootteberekeningen

Statistische effectgrootte berekeningen geven verbetering of verslechtering in de gemiddelde functionele gezondheidstoestand van een groep patiënten weer. Hierbij wordt rekening gehouden met de in een groep aanwezige verschillen (spreiding) in gezondheid. De grootte van veranderingen wordt mede bepaald door veranderingen in de binnen de groep aanwezige spreiding. Dit kan door de effectgrootte te relateren aan het maximaal te behalen effect in een patiëntengroep.

van 55 huisartsen bij 980 patiënten met niet-specifieke buikklachten werd onderzocht of de SIP afwijkende klinische effecten tussen subpopulaties betrouwbaar meet. De onderzochte groep patiënten kon worden onderverdeeld in vijf diagnostische groepen met een verschillend klinisch beloop van de klachten na respectievelijk één en zes maanden. In elk van de vijf diagnostische categorieën werden veranderingen in de functionele toestand geconstateerd overeenkomstig het klinische beloop van de klachten. Deze veranderingen in de functionele toestand waren gedeeltelijk afhankelijk van de aard van de buikklachten en van andere medische en niet-medische kenmerken (andere klachten, leeftijd).

De gevonden intra-klassecorrelaties bevestigden de reproduceerbaarheid van stabiele verschillen tussen patiënten. Tweederde van de patiënten had chronische klachten, waarbij slechts tijdelijk geringe behandelingseffecten optraden. Tevens bleek de SIP verschillen in klinische effecten tussen de groepen patiënten zeer adequaat te registreren: er werden hoge responsiviteitscorrelaties gevonden. Lage reproduceerbaarheidscoëfficiënten werden daarentegen gevonden in die categorieën waarbij grote verschillen in het klinische beloop optraden. De statistische berekeningen van de grootte van de veranderingen in de gezondheidstoestand lieten zien, dat de omvang en richtingen van de functionele veranderingen overeenkwamen met vooraf geformuleerde hypothesen over het klinisch beloop in de vijf diagnostische categorieën.

In een onderzoek van *Rutten-Mölken*²⁸ werd deze benadering eveneens toegepast. Zij vergeleek de responsiviteit en de reproduceerbaarheid van de SIP met twee ziekte-specifieke astma-vragenlijsten bij een klinische trial naar de effecten van medicamenteuze behandeling op de functionele toestand bij een groep patiënten met matig ernstige astma. De reproduceerbaarheid van de SIP was hoog. Conform de ziekte-specifieke lijsten lieten ook de SIP-gegevens een verbetering in het functioneren zien. Het instrument bleek echter on-

voldoende responsief om een verschil in behandelingseffect op de kwaliteit van leven tussen twee medicamenteuze behandelingen aan te tonen. Dit verschil in behandelingseffect kwam met de ziekte-specifieke lijsten wel duidelijk naar voren.

Ook het onderzoek van *Schuling* naar de effecten van zorg door de huisarts dan wel in het ziekenhuis op de kwaliteit van leven bij CVA-patiënten geeft aanwijzingen dat bij bepaalde ziektecategorieën een generiek instrument als de SIP onvoldoende specifiek is om kleine ziekte-specifieke behandelingseffecten te onderscheiden.²² In de periode tussen twee en zes maanden na het optreden van het cerebrovasculair accident vond *Schuling* aanwijzingen voor een kleine verbetering in de zelfredzaamheid van de patiënten, die bij analyse van de SIP-scores niet naar voren kwam.

In het onderzoek van *Schuling* zitten echter twee valkuilen.

- De eerste valkuil betreft de fasering van de meetmomenten. Hierbij gaat het erom vooraf vast te stellen op welke momenten, bij welke aspecten van de gezondheidstoestand, veranderingen in de gezondheidsuitkomsten verwacht kunnen worden. *Schuling* beschrijft dat grote verbeteringen in het algemeen functioneren na een CVA vooral optreden in de periode tussen de eerste week en de eerste twee maanden. Hij deed echter de meeste metingen met de SIP pas ná die eerste twee maanden. De veranderingen in het algemeen functioneren zijn dan dusdanig klein dat van tevoren verwacht mocht worden dat de SIP hiervoor geen geschikt instrument is.

- De tweede valkuil betreft de analyse van de scores: daarbij moet afzonderlijk aandacht worden besteed aan die aspecten waarvan verwacht mag worden dat ze door interventies van de huisarts beïnvloed worden. Bij het onderzoek van *Schuling* valt dan te denken aan aparte analyses voor categorieën als activiteiten op het gebied van de vrijetijdsbesteding, sociale interacties en het uitvoeren van bezigheden in en om het huis. Hieruit had mogelijk een andere conclusie over de responsiviteit kunnen worden getrokken dan uit de analyse over de globale indexscores.

De bevindingen in een revalidatie-onderzoek door *Mulders et al.*³⁰ kunnen voor dit laatste als illustratie dienen. De globale SIP-dimensiescores lieten geen basale veranderingen zien. Analyse van de categoriescores toonden echter aan dat functionele verbeteringen waren opgetreden in de twee gezondheidsdimensies waarbinnen op basis van de therapeutische interventies verbeteringen te verwachten waren.

Beschouwing

De Sickness Impact Profile kan een waardevolle bijdrage leveren bij onderzoek naar de effecten van het handelen van artsen op de functionele toestand van patiënten. Sinds 1987 beschikt de vakgroep huisartsgeneeskunde van de Universiteit Utrecht over een gevalideerde Nederlandse versie, die binnen klinimetrisch onderzoek kan worden gebruikt.⁴² Daarnaast heeft een onderzoeksgroep van de Rijksuniversiteit Limburg in 1994 een verkorte SIP-versie gepubliceerd. Deze versie is teruggebracht tot 68 uitspraken, waaraan de categorie- en dimensiestructuur is aangepast.⁴³ Als de duur van de afname een probleem is, zou deze verkorte versie een bruikbaar alternatief kunnen zijn.

Voor een verantwoorde toepassing van de SIP moet aandacht worden besteed aan enkele punten (zie ook tabel 3).

- Zowel patiënten als interviewers moeten nauwgezet worden geïnstrueerd over het referentiekader van de SIP. Veel gemaakte fouten betreffen: het refereren aan één specifieke klacht in plaats van de totale gezondheidstoestand, het nagaan of activiteiten op de dag van afname zijn uitgevoerd in plaats van dat deze op dat moment op de patiënt van toepassing zijn, het door de patiënt beamen van verstoringen zonder een bewuste referentie naar de gezondheidstoestand. Door tijdens de afname terug te komen op de scorereferenties wordt de betrouwbaarheid van de gegevens verhoogd, zowel bij zelfregistratie als in een interview. Daarnaast zal de betrouwbaarheid van de scores verbeteren door – met name voor oudere patiënten – een eendui-

dige omschrijving te geven van 'goede gezondheid'.

- Het afnemen van de SIP in een interview nodigt uit tot vragen om verheldering en meningen, en om het geven van ongecontroleerde uitleg. Door een gerichte training van de interviewers, met name in de vorm van proefinterviews binnen een rollenspel, wordt een hogere betrouwbaarheid van de scores bereikt.

- Ook de keuze van de registratiemethode verdient aandacht. Het variëren van de afnametechniek binnen één registratiemoment is niet wenselijk doordat niet te identificeren verschillen in 'afnameruis' tussen de scores worden geïntroduceerd. Het variëren van de afnamemethode bij herhaalde metingen is minder bezwaarlijk. De mogelijke ruis valt statistisch te determineren, terwijl daarnaast de invloed van de registratie methodes kleiner zal worden doordat de ondervraagde inmiddels bekend is met het instrument.

Bij grote onderzoeksgroepen zijn schriftelijke afnames meestal onvermijdelijk. Als de eerste afname door een interviewer wordt verricht, zal de patiënt beter in staat zijn vervolgmetingen zelfstandig in te vullen, waardoor een hogere betrouwbaarheid van de scores kan worden verkregen.

In onderzoek waarbij veel patiënten niet in staat zijn zelf de vragen te beantwoorden, zoals bij ernstig zieke patiënten, komt de generaliseerbaarheid van de bevindingen onder druk staan is. Afname bij proxies voor de hele onderzoeksgroep is dan een belangrijk alternatief. Het door elkaar heen verzamelen van gegevens van proxies en patiënten zelf geeft onbetrouwbare gegevens. Bij longitudinaal onderzoek met ontbrekende waarnemingen op de verschillende meetmomenten bieden vervangende scores een goed alternatief.⁴⁴

Literatuur

- 1 Van Crevel H, Van Gijn J. Klinimetrie: hoe gaat het met de patiënt? *Ned Tijdschr Geneesk* 1990; 134: 7-11.
- 2 Van Weel C, Meyboom-de Jong B, Van Weert H. Het functioneren van de patiënt; klinimetrie in de huisartspraktijk. *Ned*

- Tijdschr Geneesk* 1990; 134 1039-43.
- 3 Wilkin D, Hallam L, Dogget MA. Measures of need and outcome for primary health care. New York: Oxford University Press, 1992.
- 4 König-Zalm C, Furer JW, Tax B. Algemene gezondheid: profielen en indices. Assen: Van Gorcum, 1993.
- 5 Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile; development and final revision of a health status measure. *Med Care* 1981; 19: 787-805.
- 6 Bergner M, Bobbitt RA, Kressel S, et al. The sickness impact profile: conceptual formulation and methodology for the development of a health status measure. *Int J Health Serv* 1976; 6: 393-415.
- 7 Greenwald HP. The specificity of quality of life measures among the seriously ill. *Med Care* 1987; 25: 642-51.
- 8 Brooks WB, Jordan JS, Divine GW, et al. The impact of psychological factors on measurement of functional status. Assessment of the Sickness Impact Profile. *Med Care* 1990; 28: 793-804.
- 9 Sullivan M, Ahlmen M, Bjelle A. Health status assessment in rheumatoid arthritis. *J Rheumatol* 1990; 17: 439-47.
- 10 Rothmann ML, Hedrick S, Inui T. The sickness impact profile as a measure of the health status of non-cognitively impaired nursing home residents. *Med Care* 1989; 27: S157-67.
- 11 De Bruin AF, De Witte LP, Stevens J, Diederiks J. Sickness Impact Profile: the state of the art of a generic functional status measure. *Soc Sci Med* 1992; 35: 1003-14.
- 12 Jones PW and the Nedocromil Sodium Quality of Life Study Group. Quality of life, symptoms and pulmonary function in asthma: long-term treatment with nedocromil sodium examined in a controlled multi-centre trial. *Eur Resp J* 1994; 7: 55-62.
- 13 Flaherty JA, Gaviria FM, Pathak D, et al. Developing instruments for cross-cultural psychiatric research. *J Nerv Ment Dis* 1988; 176: 257-63.
- 14 Luttik A, Jacobs HM, De Witte LP. De Sickness Impact Profile: een meetinstrument, waarmee de invloed van ziekte en/of gezondheidsklachten op het dagelijks functioneren kan worden vastgesteld. Nederlandse versie. Utrecht: UH-Infom6. Vakgroep Huisartsgeneeskunde, RU Utrecht, 1987.
- 15 Jacobs HM. Health status measurement in family practice research: the sickness impact profile and its application in a follow-

- up study in patients with non-specific abdominal complaints [Dissertatie]. Utrecht: Universiteit Utrecht, 1993.
- 16 Schrier AC, Dekker FW, Kaptein AA, Dijkman JH. Quality of life in elderly patients with chronic lung disease seen in family practice. *Chest* 1990; 4: 894-9.
- 17 Berends H, Hermsen C, Jansen W, Mulder B. Een onderzoek naar de kwaliteit van leven bij patiënten met diabetes mellitus type 2 in de huisartspraktijk. Leiden; Vakgroep huisartsgeneeskunde, 1992.
- 18 Bronswijk J, Hooyen R, Mens E. Oostindier A. Kwaliteit van leven bij Hypertensie. Leiden; Vakgroep huisartsgeneeskunde, 1992.
- 19 Van Zwan AAC. Kwaliteit van leven van oudere patiënten met Cara in de huisartspraktijk. Leiden: Vakgroep huisartsgeneeskunde, 1991.
- 20 Bock de GH. Peripheral Osteoarthritis in General Practice. Zoetermeer: Proefschrift Leiden, 1994.
- 21 Dekker FW. Asthma and COPD in general practice: Studies on the quality of care. 1993.
- 22 Schuling J. Stroke patients in general practice. Impact, recovery and the GP's management in the first six months poststroke. Groningen; Styx. 1993.
- 23 Reis Miranda D, Sprangenberg JFA. Kwaliteit, doelmatigheid en organisatie van intensieve care units in Nederland. Groningen: Foundation for Research on Intensive Care in Europe, 1992.
- 24 Arendsen GH. Nierinsufficiëntie en sociaal functioneren. Utrecht: Stichting voor sociale gezondheidszorg, 1992.
- 25 Nihom J. Effecten van chronische dialyse en niertransplantatie op het zenuwstelsel [Dissertatie]. Amsterdam: Universiteit van Amsterdam, 1992.
- 26 Essink-Bot ML, Van Agt HME, Bonsel GJ. NHP of SIP: een vergelijkend onderzoek onder chronisch zieken. *Tijdschr Soc Geneesk* 1992; 70: 152-9.
- 27 De Boer JB, Van Dam, Sprangers MAG. A longitudinal study on the QL of symptomatic patients (CDC IV) in a trial of zidovudine vs viduzudine and inteleron. AIDS, in press.
- 28 Rutten-van Mólken. MPMH. Costs and effects of pharmacotherapy in asthma and COPD [Dissertatie]. Maastricht: Rijksuniversiteit Limburg, 1994.
- 29 De Witte LP, Tilli DPJ, Ticheler AJG, et al. Leven met een reumatische aandoening. Een onderzoek naar de ervaren kwaliteit van het leven bij 372 mensen met een reumatische aandoening Hoensbroek: Insti-

- tuut voor Revalidatie vraagstukken, 1989.
- 30 Mulders AHM, De Witte LP, Diederiks JPM. Evaluation of a rehabilitation after care programme for stroke patients. *J Rehab Sci* 1989; 2: 97-102.
- 31 Van Balen HGG. Mensen met een traumatisch hersenletsel: Probleeminventarisatie. Nijmegen: Sint Maartenskliniek, 1992.
- 32 De Witte LP, Philipsen H, Van de Maegdenbergh M. Stabiliteit van SIP-scores in een periode van 3 weken bij RA patiënten. Maastricht: Rijksuniversiteit Limburg, 1987.
- 33 Pollard WE, Bobbit RA, Bergner M. Examination of variable errors of measurement. *Soc Indicators Res* 1978; 5: 270-301.
- 34 Jacobs HM, Luttik A, Touw-Otten FWMM, et al. Measuring impact of sickness in patients with non-specific abdominal complaints in a Dutch family practice setting. *Med Care* 1992; 30: 244-51.
- 35 Hulsebos RG, Beltman FW, Reis Miranda D, Spangenberg JFA. Measuring quality of life with the sickness impact profile: a pilot study. *Intensive Care Med* 1991; 17: 285-8.
- 36 Pollard WE, Bobbit RA, Bergner M, et al. The Sickness impact profile: reliability of a health status measure. *Med Care* 1976; 14: 146-55.
- 37 Rothman ML, Hedrick SC, Bulcroft KA, et al. The validity of proxy generated scores as measures of patient health status. *Med Care* 1991; 29: 115-24.
- 38 Sprangers MAG, Aaronson NK. The role of Health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. *J Clin Epidemiol* 1992; 45: 743-60.
- 39 Connerly JE, Philbrick JT, Smit GR, et al. Health perceptions of primary care patients and the influence on health care utilization. *Med Care* 1989; 27: S99-109.
- 40 Epstein AM, Hall JA, Tognetti J, et al. Using proxies to evaluate quality of life. *Med Care* 1989; 27: S91-8.
- 41 Bindman AB, Keane D, Luri N. Measuring health changes among severely ill patients. The floor phenomenon. *Med Care* 1990; 28: 1143-52.
- 42 Jacobs HM, Dessens JAG. Which outcome measure? Methodological review. In: Hutchinson A, ed. Health outcomes for primary and outpatient care. Newcastle upon Tyne, in press.
- 43 De Bruin AF, Diederiks JPM, De Witte LP, Stevens FCJ. The development of a short generic version of the Sickness Impact Profile. *J Clin Epidemiol* 1994; 47: 407-18.
- 44 Jacobs HM, Touw-Otten FWMM, De Melker RA. The evaluation of changes in functional health status in patients with abdominal complaints. *J Clin Epidemiol*, in press.
- 45 Jacobs HM, Luttik A, Touw-Otten FWMM, De Melker RA. De Sickness Impact Profile; resultaten van een valideringsonderzoek van de Nederlandse versie. *Ned Tijdschr Geneeskde* 1990; 134: 1950-4.