

TNO-rapport
PG/VGZ/98.053

Levensverwachting in goede geestelijke gezondheid: Inventarisatie van potentiële instrumenten en gegevens

TNO Preventie en Gezondheid

Volksgezondheid
Gortergebouw: Wassenaarseweg 56
Postbus 2215
2301 CE Leiden

Telefoon 071 518 18 18
Fax 071 518 19 20

Datum

december 1998

Auteur(s)

Rom JM Perenboom
Loes M van Herten

TNO Preventie en Gezondheid
Gorterbibliotheek

18 MRT 1999

Postbus 2215 - 2301 CE Leiden

Het kwaliteitssysteem van
TNO Preventie en Gezondheid
voldoet aan ISO 9001.

Stamboeknummer

16618

Alle rechten voorbehouden.
Niets uit deze uitgave mag worden
vermenigvuldigd en/of openbaar
gemaakt door middel van druk, foto-
kopie, microfilm of op welke andere
wijze dan ook, zonder voorafgaande
toestemming van TNO.

Indien dit rapport in opdracht werd
uitgebracht, wordt voor de rechten en
verplichtingen van opdrachtgever en
opdrachtnemer verwezen naar de
Algemene Voorwaarden voor onder-
zoeks opdrachten aan TNO, dan wel
de betreffende terzake tussen de
partijen gesloten overeenkomst.
Het ter inzage geven van het
TNO-rapport aan direct belang-
hebbenden is toegestaan.

© 1998 TNO

Auteur

Rom JM Perenboom

ISBN-nummer

90-6743-570-8

Deze uitgave is te bestellen door het overmaken van f 21,-- (incl. BTW) op postbankrekeningnummer 99.889 ten name van TNO-PG te Leiden onder vermelding van bestelnummer *PG/VGZ/98.053*

Samenvatting

Ten behoeve van VTV 1997 is een eerste proeve van de berekening van de Levensverwachting in Goede Geestelijke Gezondheid (LGGG) gegeven. Voor deze eerste proeve is gebruik gemaakt van gegevens over de geestelijke (on)gezondheid verkregen met de Affect Balance Scale (ABS), zoals opgenomen in de CBS Continue gezondheidsenquête.

In dit rapport is geïnventariseerd welke andere instrumenten in aanmerking komen als alternatief voor de ABS, gegeven het feit dat de ABS, hoewel veel toegepast, mede door de verschillende verschijningsvormen, niet is gevalideerd.

Uitgangspunt voor deze inventarisatie is geweest dat het moest gaan om gevalideerde, internationaal toegepaste en in het Nederlands vertaalde instrumenten. In een vooronderzoek is besloten dat met name de General Health Questionnaire en in het bijzonder de twaalf item versie (GHQ-12) en de Mental Health Inventory (5-item versie, MHI-5) - onderdeel van de SF-36 - een nader onderzoek verdienen.

In dit rapport worden deze instrumenten aan de hand van internationale literatuur onderzocht op hun validiteit en betrouwbaarheid en vervolgens worden de mogelijkheden beschreven om deze instrumenten te gebruiken voor het berekenen van gezonde levensverwachtingen. Hierbij is ook gekeken naar de mogelijkheden om met behulp van deze instrumenten de ongezondheid onder te verdelen in niveaus van ernst en naar de mogelijkheden om aan deze niveaus een wegingsfactor toe te kennen waarmee een gewogen gezonde levensverwachting berekend kan worden.

In dit rapport is tevens gekeken naar de huidige beschikbaarheid van gegevensbestanden, waarin de GHQ en/of de MHI-5 zijn gebruikt. Voor de berekening van een gezonde levensverwachting hebben veel van bestanden waarin de GHQ en/of de MHI-5 zijn opgenomen het nadeel dat zij betrekking hebben op slechts een deel van de populatie. Met name het feit dat veel bestanden (zoals het NEMESIS-bestand of de MORGEN-studie) een bovengrens voor de leeftijd hebben (bij bijvoorbeeld 60 of 65 jaar) maakt ze minder geschikt voor de berekening van een gezonde levensverwachting.

Wat betreft de GHQ is het bestand van de Nationale Studie naar Ziekten en Verrichtingen in de Huisartspraktijk, uitgevoerd door het NIVEL, een optie, omdat deze studie geen leeftijdsgrenzen kent. De studie is echter uitgevoerd in 1987-1988 en daarmee niet actueel. Er zijn plannen om in 1999-2000 een nieuwe nationale studie te verrichten, waarin de GHQ-12 wederom is opgenomen (NIVEL, persoonlijke mededeling).

Voor de MHI-5 is de TNO-studie uit 1996 waarmee referentiegegevens met betrekking tot de SF-36 voor de Nederlandse populatie vanaf 16 jaar zijn verzameld een optie, omdat ook hier geen leeftijdsgrens aanwezig is.

Het is op dit moment nog niet mogelijk om op basis van de MHI-5 of de GHQ trendberekeningen te maken, omdat geen van de herhaalde metingen heeft.

Uit vergelijking van de instrumenten blijkt dat de MHI-5 en de GHQ-12 wat betreft het meten van de prevalentie van geestelijke ongezondheid op populatieniveau niet voor elkaar onderdoen. Ten opzichte van de GHQ-12 heeft de MHI-5 het voordeel dat hij deel uitmaakt van de SF-36, waarmee ook andere gezondheidsconcepten gemeten worden. Ook zijn er plannen zijn om de SF-36 standaard op te nemen in de Continue Gezondheidsenquête van het CBS.

Algehele conclusie uit deze inventarisatie is dat de MHI-5 wordt aanbevolen voor de berekening van de LGGG. Om de MHI-5 ook te gebruiken voor een naar ernstniveaus onderverdeelde Levensverwachting in Goede Geestelijke Gezondheid en een gewogen versie daarvan zijn aanvullende analyses op empirisch materiaal nodig. Het rapport sluit dan ook af met een voorstel voor nader onderzoek. Omdat ook de GHQ-12 internationaal veel wordt toegepast wordt aanbevolen om ook nader onderzoek te doen naar toepassing van dit instrument.

In een bijlage worden tenslotte (tentatieve) resultaten weergegeven van een eerste berekening van de LGGG, berekend op basis van de MHI-5.

Inhoud

Samenvatting.....	3
Inhoud	5
1 Inleiding	7
2 Instrumenten.....	9
2.1 Mental Health Inventory (MHI-5)	9
2.2 General Health Questionnaire (GHQ)	14
2.3 Affect Balance Scale	19
2.4 Vergelijking van de instrumenten vanuit de literatuur	21
2.5 Conclusies en aanbevelingen.....	22
3 Gegevensbestanden en prevalenties	24
3.1 Gegevensbestanden m.b.t. de MHI-5	24
3.2 Gegevensbestanden m.b.t. de GHQ-12.....	26
3.3 Gegevensbestanden m.b.t. de ABS.....	27
4 Conclusies en aanbevelingen	28
Literatuur.....	31
Bijlage	35

1 Inleiding

Ten behoeve van VTV-1997 is een eerste proeve gegeven van de levensverwachting in goede geestelijke gezondheid in Nederland (LGGG) (Perenboom & van de Water, 1997a; Perenboom & van de Water, 1997b). Ook voor andere landen zijn de afgelopen jaren berekeningen gemaakt van de levensverwachting in goede geestelijke gezondheid (Gispert et al., 1998; Brønnum-Hansen & Rasmussen, 1996; Mutafova et al., 1998).

Algemeen kenmerk van het idee van gezonde levensverwachting is dat de totale levensverwachting¹ met behulp van prevalentie- of incidentiegegevens over de ongezondheid wordt verdeeld in jaren in gezondheid en jaren in ongezondheid. In de klassieke benadering van de gezonde levensverwachting wordt uitsluitend een onderscheid gemaakt in gezonde en ongezonde jaren. In de VTV 1997 is deze benadering uitgebreid door het aantal jaren in ongezondheid op te splitsen in jaren in bijvoorbeeld milde ongezondheid, in matige ongezondheid en in ernstige ongezondheid. Tenslotte is het ook mogelijk om aan de verschillende niveaus van ongezondheid een wegingsfactor toe te kennen en deze mee te nemen in de berekening van het aantal gezondheidsequivalenten.. In een dergelijke gewogen benadering zou bijvoorbeeld één jaar in ernstige ongezondheid equivalent zijn aan een half jaar in goede gezondheid.

Voor de eerste Nederlandse berekeningen van de levensverwachting in goede geestelijke gezondheid (LGGG) zijn prevalentiegegevens over geestelijke ongezondheid gebruikt uit de Gezondheidsenquête van het CBS (1989-1995), waarin de 5 negatieve items van de Affect Balance Scale (ABS) zijn opgenomen (Bradburn 1969). Voor een uitgebreide beschrijving wordt verwezen naar deel III van VTV-1997 (Perenboom & van de Water, 1997a).

Hoewel de ABS in diverse onderzoeken is gebruikt om het psychologisch welbevinden van respondenten te meten - een korte inventarisatie in Medline gaf ruim 70 referenties - is dit instrument nog niet gevalideerd en wordt het instrument op vaak verschillende wijzen gebruikt. Zo kent de originele versie van de ABS 5 negatief en 5 positief geformuleerde items en per item 2 antwoordcategorieën. Het CBS heeft alleen de vijf negatief geformuleerde items in de gezondheidsenquête (sinds 1989) opgenomen en hanteert 4 antwoordcategorieën per item.

¹ het aantal jaren dat iemand van een bepaalde leeftijd in een bepaald kalenderjaar nog kan verwachten te leven

Vanwege de vele onduidelijkheden rond de ABS en omdat het instrument minder valide lijkt te zijn (Bradburn & Caplovitz, 1995), heeft het RIVM aan TNO-PG opdracht gegeven ten behoeve van VTV-2001 een inventarisatie te maken van andere instrumenten en vervolgens van gegevensbestanden die zich lenen voor de berekening van de levensverwachting in goede geestelijke gezondheid. Om te voorkomen dat de inventarisatie verandert in een mer-à-boire, zijn enkele restricties aangebracht. Zo heeft de inventarisatie van instrumenten zich (naast de ABS) beperkt tot generieke instrumenten, die internationaal veel gebruikt worden (ook voor de berekening van de LGGG), in Nederland zijn toegepast en zijn gevalideerd. Na toepassing van deze criteria heeft de inventarisatie zich beperkt tot drie generieke instrumenten, namelijk de subschaal MHI-5 van de SF-36, de GHQ-12 en de ABS.

Voor de inventarisatie van gegevensbestanden is vooral gekeken naar die bestanden die de algemene populatie betreffen, dus niet specifieke subgroepen (uitgezonderd specifieke leeftijds en/of geslachtsgroepen).

In het navolgende zal eerst ingegaan worden op de drie bovengenoemde instrumenten. In hoofdstuk 2 wordt een korte beschrijving van het instrument gegeven, met daarbij speciale aandacht voor de validiteit en betrouwbaarheid, zoals in de literatuur is gevonden. Vervolgens wordt ingegaan op de toepassingsmogelijkheden voor de berekening van een gezonde levensverwachting. Bij deze toepassingen gaat het om

- 1) de klassieke dichotome versie van de LGGG (gezond versus ongezond),
- 2) om de mogelijkheid tot indeling naar ernst van ongezondheid en
- 3) om een gewogen versie.

In hoofdstuk 3 wordt ingegaan op de geschikte en beschikbare gegevensbestanden voor de verschillende instrumenten.

In hoofdstuk 4 worden conclusies getrokken en aanbevelingen voor een praktische uitwerking gedaan.

2 Instrumenten

2.1 Mental Health Inventory (MHI-5)

Beschrijving

De MHI-5 is een van de 8 subschalen van de SF-36. Deze SF-36 (voluit geheten Medical Outcomes Study Short-Form Health Survey 36-items) is een instrument, dat is ontwikkeld om de gezondheidsstatus (vaak aangeduid met het begrip gezondheidsgerelateerde kwaliteit van leven) van personen te meten (Ware et al., 1997; Ware & Sherbourne, 1992). De SF-36 is een multi-dimensionele schaal, bestaande uit 8 subschalen die ieder een gezondheidsconcept (dimensie) meet:

- 1 beperkingen in fysieke activiteiten (10 items);
- 2 beperkingen in sociale activiteiten (2 items);
- 3 beperkingen in rol activiteiten door fysieke problemen (4 items);
- 4 lichamelijke pijn (2 items);
- 5 algemene geestelijke gezondheid (psychologische (on)welbevinden) (5 items);
- 6 beperkingen in rol activiteiten door emotionele problemen (3 items);
- 7 vitaliteit (4 items);
- 8 algemene ervaren gezondheid (5 items).

Tenslotte wordt met één item gevraagd naar veranderingen in de gezondheid.

Naast de SF-36 bestaat een nagenoeg gelijk instrument, de RAND-36. De Engelstalige items van de SF-36 zijn gelijk aan de items in de RAND-36, maar de wijze van scoring van de RAND-36 en de SF-36 verschillen iets van elkaar. In Nederland is de RAND-36 vertaald en gevalideerd door Van der Zee en Sanderman (van der Zee & Sanderman, 1993). In het kader van het IQOLA-project is de SF-36 versie in het Nederlands vertaald door Aaronson, waarbij de vertaling iets afwijkt van de vertaling van de RAND-36 (Aaronson et al., 1998; Aaronson et al., 1992).

De SF-36 is ontworpen voor zelf-rapportage en voor gebruik in zowel een klinische praktijk als in algemene gezondheidssurveys (Ware & Sherbourne, 1992). De scores van de verschillende items worden per subschaal gesommeerd. Een lage waarde (0) geeft een slechte gezondheidstoestand weer en een hoge score (100) een goede gezondheidstoestand. Het totaal resultaat op de SF-36 wordt weergegeven in een profiel van 8 scores (Ware et al., 1997).

Naast de scoring per subschaal en het daaruit resulterende profiel wordt soms ook een tweedeling gemaakt in de SF-36: een meer fysiek gericht deel (physical component summary, PCS) en een mentaal gericht deel (mental component summary, MCS) (Ware et al., 1995; Jenkinson & Layte, 1997; Ware et al., 1996; Johnson & Coons, 1998).

Een van de subschalen van de SF-36 is speciaal gericht op het meten van de geestelijke gezondheid, de Mental Health Inventory (MHI-5). De subschaal bestaat uit vijf items. Deze bovengenoemde (minimale) verschillen in vertaling zijn ook aanwezig bij de items van MHI-5 (zie tabel 1).

Tabel 1 Items van de MHI-subschaal, SF-36 versie (Aaronson et al., 1998) en RAND-36 versie (Zee & Sanderman, 1993)

	SF-36 versie	RAND-36 versie
introductie	Deze vragen gaan over hoe u zich voelt en hoe het met u ging in de afgelopen 4 weken. Wilt u a.u.b. bij elke vraag het antwoord geven dat het best benadert hoe u zich voelde. Hoe vaak gedurende de afgelopen 4 weken	Deze vragen gaan over hoe u zich de afgelopen 4 weken heeft gevoeld. Wilt u bij elke vraag het antwoord aankruisen dat het beste aansluit bij hoe u zich heeft gevoeld. Hoe vaak gedurende de afgelopen 4 weken
1	Was u erg zenuwachtig?	Voelde u zich erg zenuwachtig?
2	Zat u zo in de put dat niets u kon opvrolijken?	Zat u zo in de put dat niets u kon opvrolijken?
3*	Voelde u zich rustig en tevreden?	Voelde u zich kalm en rustig?
4	Voelde u zich somber en neerslachtig?	Voelde u zich neerslachtig en somber?
5*	Was u een gelukkig mens?	Voelde u zich gelukkig?
respons mogelijk- heden	1 Altijd 2 Meestal 3 Vaak 4 Soms 5 Zelden 6 Nooit	1 Voortdurend 2 Meestal 3 Vaak 4 Soms 5 Zelden 6 Nooit

* De scores voor deze items worden gehercodeerd (1 = 6, 2=5 et cetera)

Met deze vijf items worden diverse mentale gezondheids dimensies bestreken, namelijk angst, depressie, controleverlies over gedrag of emoties en psychologisch welbevinden (Ware & Sherbourne, 1992). Per item zijn er zes antwoordcategorieën aanwezig met een score van 1 tot en met 6. Hercodering (waar noodzakelijk) en sommering leidt tot 26 niveaus van geestelijke gezondheid (range 5 t/m 30), waarbij een hoge score een goede gezondheid aangeeft en een lage score een slechte. Voor de weergave in het SF-36 profiel wordt de oorspronkelijke scorereange omgewerkt tot een range van 0 tot 100, door iedere score volgens een specifiek algoritme te bewerken (zie Ware et al., 1997). Een lage waarde op deze schaal geeft aan dat iemand altijd gevoelens van nervositeit en

depressie heeft, terwijl een hoge score op deze range aangeeft dat iemand zich altijd vredig, gelukkig en kalm voelt (Ware et al., 1997).

Volgens de ontwerpers is de MHI-subschaal ontwikkeld als een bipolaire schaal, die een grote range van gezondheidsstatussen meet. Een lage score geeft een slechte mentale gezondheid weer, een score in het middengebied (rond de score 50) geeft aan dat er geen mentale problemen zijn, terwijl een hoge score (100) een goede tot zeer goede mentale gezondheid weerspiegelt. Dit in tegenstelling tot SF-36-subschalen als fysiek functioneren of pijn, waar de hoogste score (100) de afwezigheid van beperkingen respectievelijk pijn aangeven (Ware et al., 1997). In enkele Nederlandse studies blijkt de gemiddelde score op de MHI-5 in een normale populatie ruim boven de score van 50 te liggen. Zo ligt de gemiddelde score in de Morgenstudie rond de 80 (Edlinger et al., 1998) en in de TNO-referentie-studie op circa 76 (Aaronson, concept). Voor verdere gegevens over deze studies wordt verwezen naar hoofdstuk 3.1.

Validiteit en betrouwbaarheid

Om de validiteit en de waarde van de MHI-5 voor het meten van geestelijke gezondheid en ongezondheid te bepalen is een aantal argumenten in de literatuur gevonden. Zo vormen de items van de MHI-5 een weloverwogen selectie van de items uit de 38-item mental health inventory (MHI-38) van de Medical Outcomes Study (MOS). De MHI-38 is volgens Ware et al. (1997) via veel studies gevalideerd voor het meten van de prevalentie van geestelijke ongezondheid.

De interne consistentie, gemeten met Cronbach's α , ligt voor de MHI-5 bij de meeste studies waarin hierover gerapporteerd wordt boven de 0.80 (Ware et al., 1997; Brazier et al., 1995; Stewart & Ware, 1993).

Het discriminerend vermogen van de MHI-5 (het vermogen onderscheid te maken tussen ongezonden en gezonden) wordt als goed aangegeven (Stewart & Ware, 1993). Zo blijken patiënten in de geestelijke gezondheidszorg slechts minimaal voor te komen in de hoogste subschaalscores (80-100). Afgezet tegen een vijftal problemen (ontevredenheid met het leven, depressieve symptomen, diagnose depressie, zelfmoord-ideeën, en gebruik van geestelijke gezondheidszorg) blijkt dat er een lineair verband is tussen de ernst van het probleem en de gemiddelde score op de MHI-5 (Ware et al., 1997).

Om weer te geven wat een bepaalde score op de MHI-5 schaal betekent, is in de Medical Outcomes Study getracht een interpretatie te vinden voor de (verschillen in) scores. Zo blijkt het verschil tussen

de score 20 en de score 40 gelijk te staan aan 15 procentpunten verschil in de waarschijnlijkheid van een klinische diagnose voor depressie.(Ware et al., 1997).

Gebruik ten behoeve van de LGGG

Voor de berekening van gezonde levensjaren worden gewoonlijk één of meer afkappunten gebruikt om diverse niveaus van (on)gezondheid te bepalen (Perenboom et al., 1997a). Bij gebruik van één afkappunt (gezond - ongezond) is er sprake van de klassieke dichotome versie, bij gebruik van meer afkappunten is er sprake van een onderverdeling naar ernst. Deze onderverdelingen naar ernst kunnen via een wegingsfactor ook worden omgerekend naar een gezondheidsequivalent.

Voor de dichotome benadering ligt het voor de hand om een afkappunt te definiëren in het midden van de scorereange, omdat er op dat niveau volgens Ware et al. sprake is van afwezigheid van geestelijke problemen of ongezondheid (Ware et al., 1997). Brønnum-Hansen en Rasmussen hebben bij tentatieve berekeningen van de LGGG voor Denemarken, conform een studie van Berwick et al (1991), iedereen met een score van 52 en lager (op de 100 punt schaal) als ongezond beoordeeld (Brønnum-Hansen & Rasmussen, 1996). In de studie van Berwick et al. (1991) is een vergelijking gemaakt van diverse screener instrumenten, waaronder de MHI-5. Wanneer op basis van een Diagnostic Interview Schedule 96 op de 1000 personen een majeure depressie hebben worden er door de MHI-5 (bij een afkappunt tussen de scores 52 en 53) 46 ontdekt, terwijl de uitslag voor 43 personen vals positief blijkt te zijn.

Voor een naar ernst uitgesplitste LGGG dienen meerdere afkappunten te worden gedefinieerd. Net als bij de opdeling van de fysieke beperkingen naar ernstniveaus zijn er ook hier twee basisprincipes waarop deze afkappunten bepaald kunnen worden, namelijk

- 1) op basis van de individuele items
- 2) op basis van de subschaalscore

Beide basisprincipes kunnen vervolgens op verschillende wijzen verder worden uitgewerkt:

ad 1) Indien een keuze gemaakt wordt voor een indelingsprincipe op basis van de individuele items dient onderzocht te worden hoe de ernst bepaald wordt:

- 1a) Zo kan kan op basis van het aantal ongezond gescoorde items een indeling naar ernst gemaakt worden: bijvoorbeeld één ongezond item geeft milde ongezondheid weer, twee ongezonde items matige en drie of meer ernstige ongezondheid. Voor een gewogen gezonde levensverwachting

dienen aan deze niveaus vervolgens wegingsfactoren toegekend te worden. Dit kan wellicht op basis van een deskundigen-oordeel, zoals in het kader van het project Wegingsfactoren voor ziekten is gebeurd (Stouthard et al., 1997).

1b) Ook is het mogelijk aan de verschillende individuele items ernstscores (wegingsfactoren) mee te geven, hoewel in de literatuur hierover geen gegevens bekend zijn. Ware c.s. gaan er tot op heden vanuit dat de items ten opzichte van elkaar gelijkwaardig zijn (Kosinsky, persoonlijke mededeling). Indien ernstscores toegekend kunnen worden, bijvoorbeeld via een deskundigen-oordeel (zie hierboven), kan hiermee zowel een indeling naar ernstniveaus gemaakt worden als een gewogen gezonde levensverwachting berekend worden.

ad 2) Op subschaalniveau (range 0-100) kan bij ernstniveaus gedacht worden aan ranges van scores. Verstandig lijkt het in ieder geval één van de afkappunten bij de score 52 te leggen, omdat bij een score van 52 of meer geen beperkingen of ongezondheid aanwezig is en omdat daarmee een vergelijking met bijvoorbeeld Denemarken gelegd kan worden (Brønnum-Hansen & Rasmussen, 1996). Scores van 53 en hoger kunnen als gezond worden beoordeeld, de scorereange van 52 en lager kan verdeeld worden in een aantal niveaus van ongezondheid. Daarbij kan gedacht worden aan bijvoorbeeld de volgende ranges: 0-16 ernstig ongezond, 17-36 matig ongezond en 37-52 mild ongezond.

Voor de gewogen benadering is indeling van de scorereange in een aantal niveaus niet per se noodzakelijk. De verschillende scores van 0 tot 100 kunnen in principe als weegfactor in de berekening van de gewogen gezonde levensverwachting ingebracht worden. Maar om de vergelijking met de in ernstniveaus uitgesplitste gezonde levensverwachting mogelijk te maken kan beter gekozen worden voor een indeling van de scorereange onder de waarde 53 op eenzelfde wijze als boven beschreven. Vervolgens kan aan iedere niveau een wegingsfactor toegekend worden. Voor de bepaling van deze wegingsfactor zijn er verschillende opties, bijvoorbeeld met behulp van een deskundigenoordeel of op basis van empirisch materiaal. Hierop wordt in hoofdstuk 4 bij de aanbevelingen nader ingegaan.

2.2 General Health Questionnaire (GHQ)

Beschrijving

De General Health Questionnaire (GHQ) is ontworpen als een door de respondent zelf in te vullen screeningsinstrument, bedoeld om (niet psychotische) psychiatrische ziekte te detecteren (Goldberg & Williams, 1988; Goldberg, 1995). De GHQ is ontwikkeld vanuit een set van 140 vragen, verdeeld over 4 domeinen van geestelijke ongezondheid: depressie, angst, objectief observeerbaar gedrag (inclusief 'social impairment' en 'social inadequacy') en hypochondriasis.

Op basis van discriminerend vermogen en principale componenten analyse zijn 60 vragen geselecteerd die tezamen de GHQ vormen. Vanuit deze 60 vragen zijn vervolgens geneste subsets gevormd van 30, 20 en 12 vragen, waarbij telkens de sterkste items zijn gekozen. De GHQ-12 maakt dus deel uit van de GHQ-20, de GHQ-20 van de GHQ-30 en de GHQ-30 van de oorspronkelijke GHQ-60. Daarnaast bestaat nog een aparte subset, de GHQ-28, welke niet in deze geneste versies past. De GHQ-12 is een van de meest toegepaste versies van de GHQ-familie. In tabel 2 staan de vragen van de 12-item versie van de GHQ conform de Nederlandse handleiding (Koeter & Ormel, 1991)

Het belangrijkste kenmerk van de GHQ is, dat het zich richt op veranderingen in het psychisch functioneren, hetgeen zich uit in de antwoordcategorieën, die in termen zijn gesteld van 'meer of minder dan gebruikelijk' (Goldberg, 1995). Tegelijkertijd vormt dit ook een van de belangrijkste kritiekpunten, namelijk dat op negatief geformuleerde vragen een antwoord als 'niet meer dan gewoonlijk' niet per se wijst op 'normaliteit' (Finlay-Jones & Murphey, 1979; Goodchild & Duncan-Jones, 1985). Om aan deze kritiek tegemoet te komen zijn twee additionele vragen ontwikkeld, waarmee chronische psychiatrische ziekten opgespoord kunnen worden. Deze vragen hebben betrekking op het gebruik van kalmerende middelen en of respondenten denken dat ze psychische klachten hebben.

Tabel 2 Items van de Nederlandse versie van de GHQ-12 (Koeter & Ormel, 1991)

item	antwoordcategorieën
Bent u de laatste tijd door zorgen veel slaap tekort gekomen?	helemaal niet, niet meer dan gewoonlijk, wat meer dan gewoonlijk, veel meer dan gewoonlijk
Heeft u de laatste tijd het gevoel gehad dat u voortdurend onder druk stond?	helemaal niet, niet meer dan gewoonlijk, wat meer dan gewoonlijk, veel meer dan gewoonlijk
Heeft u zich de laatste tijd kunnen concentreren op uw bezigheden?	beter dan gewoonlijk, net zo goed als gewoonlijk, slechter dan gewoonlijk, veel slechter dan gewoonlijk
Heeft u de laatste tijd het gevoel gehad zinvol bezig te zijn?	zinvoller dan gewoonlijk, net zo zinvol als gewoonlijk, minder zinvol dan gewoonlijk, veel minder zinvol dan gewoonlijk
Bent u de laatste tijd in staat geweest uw problemen onder ogen te zien?	beter dan gewoonlijk, net zo goed als gewoonlijk, minder goed in staat dan gewoonlijk, veel minder goed in staat dan gewoonlijk
Voelde u zich de laatste tijd in staat om beslissingen (over dingen) te nemen?	beter in staat dan gewoonlijk, net zo goed in staat als gewoonlijk, wat minder goed in staat dan gewoonlijk, veel minder goed in staat dan gewoonlijk
Heeft u de laatste tijd het gevoel gehad dat u uw moeilijkheden niet de baas kon?	nee, dat gevoel helemaal niet, niet minder de baas dan gewoonlijk, wat minder de baas dan gewoonlijk, veel minder de baas dan gewoonlijk
Heeft u zich de laatste tijd alles bij elkaar redelijk gelukkig gevoeld?	gelukkiger dan gewoonlijk, even gelukkig als gewoonlijk, minder gelukkig dan gewoonlijk, veel minder gelukkig dan gewoonlijk
Heeft u de laatste tijd plezier kunnen beleven aan uw gewone, dagelijkse bezigheden?	meer dan gewoonlijk, even veel als gewoonlijk, wat minder dan gewoonlijk, veel minder dan gewoonlijk
Heeft u zich de laatste tijd ongelukkig en neerslachtig gevoeld?	helemaal niet, niet meer dan gewoonlijk, wat meer dan gewoonlijk, veel meer dan gewoonlijk
Bent u de laatste tijd het vertrouwen in uzelf kwijtgeraakt?	helemaal niet, niet meer dan gewoonlijk, wat meer dan gewoonlijk, veel meer dan gewoonlijk
Heeft u zich de laatste tijd als een waardeloos iemand beschouwd?	helemaal niet, niet meer dan gewoonlijk, wat meer dan gewoonlijk, veel meer dan gewoonlijk

Tweede kenmerk van de GHQ is de specifieke wijze van scoren. Per item zijn er vier responscategorieën aanwezig die van 1 tot en met 4 gescoord kunnen worden. Voor de berekening van de GHQ score worden eerst enkele items omgecodeerd, zodanig dat een score van 3 of 4 'ongezondheid' weergeeft. De antwoordcategorieën worden gedichotomiseerd (de scores 1 en 2 krijgen een waarde 0, terwijl de scores 3 en 4 een waarde 1 krijgen), waarna een somscore wordt berekend, waarbij een hoge schaalscore een staat van ongezondheid aangeeft en een lage schaalscore een staat van gezondheid.

Daarna wordt er een afkappunt gelegd om potentiële cases te detecteren. Er bestaan volgens Koeter en Ormel (1991) geen algemeen geaccepteerde drempelwaarden, maar geadviseerd wordt om aan te sluiten bij de modale waarden die Goldberg en Williams (1997) noemen. Voor de GHQ-60 ligt dit

afkappunt tussen 11 en 12, voor de GHQ-30 tussen 3 en 4 en voor de GHQ-28 tussen 4 en 5. Voor de GHQ-12 is nog geen duidelijke keuze gemaakt of dat afkappunt zal liggen tussen de score 1 en 2 of tussen de score 2 en 3. Koeter en Ormel adviseren om in de algemene populatie voor de GHQ-12 een afkappunt tussen de scores 1 en 2 te nemen (Koeter & Ormel, 1991).

Validiteit en betrouwbaarheid

Volgens Koeter en Ormel (1991) wordt de GHQ op drie manieren gebruikt:

- 1) om de ernst van de niet-psychotische psychopathologie te schatten;
- 2) als screeningsvragenlijst voor de detectie van mogelijke psychiatrische gevallen (bepaald op basis van het afkappunt), waarbij vervolgens een psychiatrisch interview wordt afgenomen om de aard van de aandoening te bepalen;
- 3) om de prevalentie van psychiatrische 'cases' in de populatie te schatten.

Ad 1) Om de ernst van de niet psychotische psychopathologie in te schatten dient het instrument betrouwbaar en valide te zijn, afgemeten aan een 'gouden' standaard. In divers onderzoek is de GHQ (in al zijn gedaanten) op validiteit en betrouwbaarheid beproefd. Uit in Nederland uitgevoerd onderzoek blijken de Cronbachs α 's van de verschillende versies meestal ruim boven de .85 te liggen, zowel in de algemene populatie als onder specifieke groepen (specialistische settings, eerste lijns gezondheidszorg) (Koeter & Ormel, 1991).

De correlatie tussen de GHQ en psychiatrische interviews varieert met zowel de versie van de GHQ als het type psychiatrisch interview. De correlatie van de GHQ-12 met bijvoorbeeld de Present-State Examination (PSE) is .53, terwijl die tussen de GHQ-12 en de Clinical Interview Schedule (CIS) .70 is (Goldberg & Williams, 1988).

Uitgaande van een één-dimensionele benadering van psychopathologie kan de GHQ ook ingezet worden om veranderingen in de ernst van psychopathologie te meten, maar dan dient rekening gehouden te worden met een hertest-effect. Hieronder wordt het volgens Koeter en Ormel nog onvoldoende verklaarde fenomeen verstaan dat de score bij een tweede afname van de GHQ is veranderd, terwijl hun gezondheidstoestand volgens bijvoorbeeld de PSE niet is veranderd. Via gebruik van een controlegroep is voor dit hertesteffect te controleren (Koeter & Ormel, 1991).

Ad 2) Indien de GHQ als screeningsinstrument (case-finding) wordt gebruikt, spelen sensitiviteit (terecht een echte 'case' herkennen) en specificiteit (een niet echte case ook als niet echt herkennen)

een belangrijke rol. De Engelse versie van de GHQ kent een sensitiviteit die varieert tussen de .74 (voor de GHQ-30) en de .89 (voor de GHQ-12) en een specificiteit van .80 (GHQ-12) tot .87 (GHQ-60) (Goldberg & Williams, 1988).

In Nederland liggen de sensitiviteit en specificiteit wat lager dan in de Engelse studies gemeten. Dit kan mede samenhangen met het feit dat voor de bepaling van de sensitiviteit en specificiteit in Engeland studies samen zijn genomen met verschillende afkappunten (Koeter & Ormel, 1991).

In Nederland loopt de sensitiviteit van de GHQ-12 in de diverse studies uiteen van .61 tot .94, terwijl de specificiteit uiteenloopt van .60 tot .78. Deze verschillen in de diverse Nederlandse studies kunnen wellicht voor een groot deel verklaard worden door de verschillende steekproeven (huisartspopulaties en algemene populaties), en steekproeftrekkingen (proportioneel en non-proportioneel) die in de behandelde studies zijn gehanteerd (Koeter & Ormel, 1991).

De verschillen die in de Engelse studies aangetroffen worden in sensitiviteit en specificiteit kennen meerdere oorzaken: ze zijn berekend op basis van diverse onderzoeken, en niet altijd is hetzelfde afkappunt gehanteerd. De plaats van het afkappunt is echter in hoge mate bepalend voor zowel de sensitiviteit en specificiteit. Daarnaast zijn de sensitiviteit en de specificiteit in deze studies geschat ten opzichte van verschillende psychiatrische interview methoden.

Bovendien is de GHQ-12 leeftijdsgevoelig, dat wil zeggen dat ouderen (ten onrechte) meer kans maken om als 'case' te worden gerekend dan jongeren. De reden is dat met de GHQ-12 ook somatische problemen gemeten kunnen worden, die zich bij ouderen eerder voordoen dan bij jongeren (bijvoorbeeld slaapproblemen). Ter correctie wordt voor ouderen vaak een ander (hoger) afkappunt te worden gehanteerd. Ook bij een onderzoekspopulatie met veel chronisch zieken wordt het afkappunt van de GHQ-12 vaak hoger gelegd (Koeter & Ormel, 1991).

Ad 3) Tenslotte is de GHQ te gebruiken om de prevalentie van psychiatrische morbiditeit in een (sub)populatie te berekenen. Om de prevalentie te bepalen is het zinvol om onderscheid te maken in de wijze van steekproeftrekking, namelijk een tweetrapssteekproef en een ééntrapssteekproef. In de tweetrapssteekproef wordt bij een selectie van de steekproef een psychiatrisch interview afgenomen.

De ééntrapssteekproef betreft een representatieve steekproef uit de populatie waarbij men op basis van de GHQ scoreverdeling een uitspraak wil doen over de prevalentie. Binnen het design van een eenetrapssteekproef zijn er drie mogelijkheden om tot een schatting van de prevalentie te komen (Koeter & Ormel, 1991):

- a) de proportie GHQ-cases (boven het afkappunt) wordt als schatter van de prevalentie gebruikt. Als de sensitiviteit en de specificiteit onder de 1 liggen, is de schatting onzuiver en dient er eigenlijk een correctie plaats te vinden, zoals beschreven onder b.
- b) met behulp van de bekende sensitiviteit en specificiteit wordt de schatting uit a. gecorrigeerd volgens de volgende formule:

$$\text{prevalentie} = (T + Sp - 1) / (Se + Sp - 1),$$

waarin T = proportie GHQ-cases in steekproef, Sp = specificiteit, Se = sensitiviteit.

Als de specificiteit en sensitiviteit exact bekend zijn, is de schatting zuiver. Indien ook de specificiteit en sensitiviteit geschat zijn, is de schatting van de prevalentie vertekend, maar in mindere mate dan de schatting uit a.

- c) Tenslotte is het mogelijk om per GHQ-score de kans om een echte case te zijn (volgens een psychiatrisch interview) te bepalen via een logistische functie (Koeter & Ormel, 1991).

Gebruik ten behoeve van de LGGG

Het gebruik van de GHQ als hierboven omschreven bij punt 3) is de toepassing die voor de LGGG van belang is. Om de prevalentie van gezondheid en ongezondheid te meten, zijn de onder 3a, 3b en 3c genoemde manieren allemaal mogelijk. Zo is in Bulgarije de LGGG berekend met behulp van de GHQ-12, conform methode 3a, met als afkappunt 2-3, waarbij de prevalentie van ongezondheid niet is gecorrigeerd voor sensitiviteit en specificiteit (Mutafova et al., 1996). Gispert daarentegen heeft op basis van een logistische regressie (methode 3c) de kansen per GHQ-12-score om een echte case te zijn berekend. De hieruit berekende prevalentie van geestelijke ongezondheid in Catalonië (Spanje) is vervolgens toegepast in de berekening van de Levensverwachting in Goede Geestelijke Gezondheid (Gispert et al., 1998; Gispert et al., 1996).

Zoals boven omschreven kan de GHQ ook gebruikt worden om veranderingen in de ernst van de psychopathologie te meten (Koeter & Ormel, 1991). Hoewel het hier uitsluitend om veranderingen lijkt te gaan, biedt dit wellicht mogelijkheden om ook een indeling naar ernstniveaus aan te brengen. Hierover zijn geen gegevens in de literatuur gevonden, maar gedacht kan worden aan het gebruik van een aantal afkappunten, bijvoorbeeld tussen de scores 2 en 3, tussen 3 en 4 en tussen 4 en 5. Voor een gewogen benadering zal dan aan de verschillende onderscheiden ernstniveaus een

wegingsfactor toegekend moeten worden. Wellicht kan dit op basis van de kansen die - op basis van een psychiatrisch interview - toegekend kunnen worden aan de verschillende GHQ-scores.

2.3 Affect Balance Scale

Beschrijving

De Affect Balance Scale (ABS) is in de zestiger jaren ontwikkeld om het alledaags psychisch onwelbevinden te meten, zoals gerapporteerd door de respondent zelf (Bradburn, 1969; Bradburn & Caplovitz, 1995).

De ABS bestaat uit tien vragen, waarvan de helft betrekking heeft op prettige gevoelens en de andere helft op onprettige gevoelens. De vijf vragen naar positieve gevoelens worden vaak aangeduid als de Positive Affect Scale (PAS), die naar de negatieve gevoelens als de Negative Affect Scale (NAS). Ook andere afkortingen worden wel gebruikt, bijvoorbeeld AMIN (voor de NAS) (Heydendael et al., 1986). In tabel 3 worden de oorspronkelijke Amerikaanse items vermeld en de Nederlandse versie zoals toegepast in de CBS Continue Gezondheidsenquête. In de oorspronkelijke versie van 1965 waren er twee antwoordcategorieën, yes en no. Na het antwoord 'yes' werd nog gevraagd naar de frequentie (once, several times, often) (Bradburn & Caplovitz, 1995). In de versie van 1969 is de vervolgvraag naar de frequentie vervallen. Omdat de ABS en de wijzigingen daarop niet echt goed onderbouwd zijn, zijn verschillende varianten in omloop gekomen, zoals al uit de CBS-versie blijkt.

Tabel 3 Items van de Affect Balance Scale (Versie Bradburn 1969 en Versie CBS-continue Gezondheidsenquête)

	Bradburn 1969- versie	CBS-versie
	During the past few weeks did you ever feel	
positive feelings		
1	pleased about having accomplished something?	-
2	that things were going your way?	-
3	proud that someone complimented you on something you had done?	-
4	particularly excited or interested in something?	-
5	on top of the world?	-

negative feelings		
1	so restless that you couldn't sit long in a chair?	Heeft u zich de afgelopen weken wel eens zo rusteloos gevoeld dat u bij wijze van spreken niet stil kon blijven zitten?
2	bored?	Heeft u zich de afgelopen weken wel eens erg verveeld?
3	depressed or very unhappy?	Heeft u zich de afgelopen weken wel eens erg tekeer geslagen gevoeld of ergens over in de put gezeten?
4	very lonely or remote from other people?	Heeft u zich de afgelopen weken wel eens erg eenzaam en verlaten gevoeld?
5	upset because someone criticized you?	Heeft u zich de afgelopen weken wel eens van streek gevoeld omdat iemand een aanmerking op u maakte?

Validiteit en betrouwbaarheid

Uit een Medline-search blijkt dat het instrument in een groot aantal studies wordt gebruikt om psychologisch (on)welbevinden en de kwaliteit van leven te meten, zowel in studies waar het gaat om het meten van effecten van medische interventies als in populatiestudies waarin de prevalentie van psychisch (on)welbevinden geschat wordt. Ondanks het feit dat de ABS (in welke gedaante dan ook) vaak gebruikt wordt, worden de betrouwbaarheid en de validiteit van de ABS regelmatig ter discussie gesteld (Bradburn & Caplovitz, 1995; Kempen & Ormel, 1998; Himmelfarb & Murrell, 1998). Deze vraagtekens hebben met name betrekking op de positieve subschaal (PAS). De NAS blijkt in onderzoek redelijk te correleren met andere indicatoren voor geestelijk onwelbevinden (Bradburn & Caplovitz, 1995). De relatie tussen een score op de ABS en de aan- of afwezigheid van psychiatrische diagnoses is echter niet vastgesteld.

Sinds 1989 zijn de negatieve items (de NAS) opgenomen in de Gezondheidsenquête van het CBS (Cronbachs $\alpha = .79$, voor het bestand 1995). Hiermee is het mogelijk om de prevalentie van geestelijk (on)welbevinden te meten, zij het dat niet exact aangegeven is welke antwoorden (scores) een toestand van geestelijk welbevinden of van onwelbevinden aanduiden.

Daarnaast is in Nederland de ABS toegepast in onderzoek bij specifieke groepen (Van der Does et al., 1998; Kempen & Ormel, 1998).

Gebruik ten behoeve van de LGGG

Met behulp van de NAS-gegevens uit de CBS Continue Gezondheidsenquête is in Nederland de LGGG berekend volgens de klassieke dichotome manier (Perenboom & van de Water, 1997a; Perenboom & van de Water, 1997b). Hierbij is iemand als geestelijk ongezond beoordeeld wanneer die minimaal 1 van de vijf items beantwoordt met 'vaak' of 'erg vaak'.

Het is ook mogelijk met behulp van de vier antwoordcategorieën een somscore te berekenen, waarbij per item het antwoord 'helemaal niet' een waarde 0 krijgt en het antwoord 'erg vaak' een waarde 3. Dit leidt tot een somscorerange van 0 tot en met 15. Heydendael et al. beschouwen een dergelijke somscore iedereen met een score van 3 of hoger als geestelijk ongezond (Heydendael et al., 1986). Zowel op basis van deze somscore als op basis van de voor VTV-1997 gebruikte wijze kan een verdere onderverdeling naar ernstniveau's worden aangebracht, alhoewel in de literatuur geen gegevens gevonden zijn over de relatie tussen ernst van de geestelijke ongezondheid (psychiatrische diagnoses) en scoring op de NAS.

2.4 Vergelijking van de instrumenten vanuit de literatuur

Uit een search in Medline is gebleken dat slechts in enkele studies de MHI-5, de GHQ (-12 en -30), en/of de NAS met elkaar zijn vergeleken op met name hun screenend vermogen, en dan in het bijzonder hoe zij ten opzichte van elkaar (en eventueel een 'gouden' standaard als een psychiatrisch interview) presteren. Zo vergelijken Heydendaal et al de GHQ-30, de GHQ-12 en de NAS (bij hen AMIN geheten) (Heydendaal et al., 1986). De specificiteit en sensitiviteit van deze verschillende instrumenten werden vastgesteld ten opzichte van de Present State Examination (PSE). Tabel 4 geeft de resultaten weer voor deze instrumenten. Hierbij zijn drempelwaarden gekozen voor de GHQ-12 en de NAS, die zo vergelijkbaar mogelijk percentages positieven geven als het afkappunt 3-4 voor de GHQ-30.

Tabel 4 Vergelijking GHQ-30, GHQ-12 en NAS (AMIN) (Heydendaal et al., 1986)

	drempelwaarde	% positief	sensitiviteit	specificiteit
GHQ-30	4	28,5	.74	.75
GHQ-12	2	25	.68	.71
NAS	3	26,8	.69	.77

Uit deze tabel blijkt dat de NAS en de GHQ-12 in sensitiviteit niet veel van elkaar verschillen. Tevens komt uit de studie van Heydendaal et al. naar voren dat de NAS circa 84% van de door de GHQ-30 als terecht ontdekte gevallen ook als zodanig classificeert (Heydendaal et al., 1986).

In een studie van McCabe et al. worden de prestaties van de GHQ-12 vergeleken met die van de MHI-5 (McCabe et al., 1998). Hierbij dient opgemerkt te worden dat de wijze van scoring van de GHQ-12 niet conform de officiële scoringsmethodiek is verlopen. McCabe et al. gebruiken voor de berekening van de somscore over de GHQ niet de gedichotomiseerde waarden, maar de oorspronkelijke itemscores (in hun studie uiteenlopend van 0 tot en met 3). Deze itemscores worden vervolgens over de 12 items opgeteld tot een somscore, die een range kent van 0 tot en met 36, waarbij 36 de slechtste gezondheid weerspiegelt. Uit de vergelijking van de instrumenten komt als belangrijkste conclusie naar voren dat de GHQ-12 en de MHI-5 vergelijkbare psychometrische prestaties hebben. Beide kunnen gebruikt worden om de mentale gezondheid in de populatie te meten. Wel blijkt dat de GHQ-12 positief gerelateerd is aan leeftijd. Hoe hoger de leeftijd, hoe groter volgens de GHQ-12 de kans is op geestelijke ongezondheid. De MHI-5 geeft deze relatie niet aan.

2.5 Conclusies en aanbevelingen

Om een beeld te geven van de prevalentie van geestelijke ongezondheid doen de drie hierboven vermelde instrumenten niet veel voor elkaar onder. Bij geen der instrumenten zijn in de literatuur mogelijkheden aangetroffen om onderscheid naar ernst van ongezondheid te kunnen maken, maar de GHQ-12 en de MHI-5 kunnen gebruikt worden om veranderingen in de geestelijke (on)gezondheid vast te leggen, hetgeen wellicht de mogelijkheid biedt voor een onderverdeling naar ernst.

De verschillende instrumenten hebben hun eigen sterke en zwakke punten.

De NAS is weliswaar nooit echt gevalideerd, maar lijkt in vergelijking met de GHQ-12 redelijk te presteren. Het voordeel van gebruik van de NAS voor de berekening van de LGGG is, dat deze sinds 1989 door het CBS is opgenomen in de gezondheidsenquête. Hierdoor zijn vanaf 1989 gegevens bekend van de zelfstandig wonende populatie boven de 16 jaar en kan de LGGG vanaf die leeftijd berekend worden.² Nadeel van de NAS is dat (mede door het grote aantal varianten dat internationaal de ronde doet) er tot op heden geen internationale vergelijkingen mogelijk zijn wat betreft prevalentie van geestelijke ongezondheid, en daarmee geen vergelijking van de LGGG.

² Overigens zijn nagenoeg geen referenties gevonden waarbij een van de drie instrumenten is toegepast op personen jonger dan 16 jaar. Een uitzondering daarop is een studie, waarin de ABS ook is afgenomen bij personen van 15 jaar (Stacey & Gatz, 1991).

De GHQ-12 en de MHI-5 hebben als voordeel dat zij in diverse studies uitvoerig zijn getest op validiteit en betrouwbaarheid en internationaal zeer veel toegepast worden. Indien de wijze van scoren en het bepalen van afkappunten internationaal op dezelfde wijze geschiedt, biedt dit de mogelijkheid om internationale vergelijkingen te maken.

De GHQ-12 heeft daarbij het nadeel dat het afkappunt niet altijd op dezelfde plaats wordt gelegd, bijvoorbeeld vanwege mogelijke leeftijdseffecten of de aanwezigheid van chronische ziekten. Hierdoor is het niet mogelijk om prevalentiegegevens, gebaseerd op de GHQ-12 zonder meer over te nemen voor toepassing in de LGGG.

De MHI-5 heeft daarbij ten opzichte van de GHQ-12 als voordeel dat het instrument kort is (5 items), waardoor gebruik in vragenlijst onderzoek wellicht op minder bezwaar zal stuiten dan langere instrumenten. Bovendien maakt het deel uit van een groter instrument (SF-36), dat acht gezondheidsdomeinen meet. Gebruik van dit gehele instrument zou de mogelijkheid kunnen bieden om niet alleen de LGGG te berekenen, maar ook andere typen gezonde levensverwachting en wellicht een soort overkoepelend GLV, gebaseerd op de resultaten van de complete SF-36. Voor toekomstige toepassingen en voor internationale vergelijking lijkt gebruikmaking van de MHI-5 als deel van de SF-36 daarom het meest belovend.

3 Gegevensbestanden en prevalenties

3.1 Gegevensbestanden m.b.t. de MHI-5

TNO-PG heeft in 1996 een survey gehouden in een steekproef uit de Nederlandse bevolking van 16 jaar en ouder (N=2800, respons 63% = 1771), waarbij het doel voornamelijk was het verkrijgen van zogenaamde nationale referentiewaarden voor de Nederlandse bevolking (Aaronson et al., 1998). De steekproef betrof een random steekproef uit het nationale telefoonboek. Hierin zijn mannen ietwat oververtegenwoordigd en personen tussen 16 en 25 jaar ondervertegenwoordigd. Via de uitnodigingsbrief is getracht hiervoor te corrigeren, door expliciet te vragen naar invulling door iemand van het vrouwelijk geslacht en/of van de leeftijdsgroep 16 tot 25 jaar (indien aanwezig). Uiteindelijk kende de responsgroep een wat hoger percentage jonge mannen en gehuwden dan de totale bevolking, hetgeen tot enige voorzichtigheid wat betreft de representativiteit noopt. In tabel 5 geven we de score verdeling voor de MHI weer voor mannen en vrouwen. De scores zijn gegroepeerd conform de in hoofdstuk 2.1 aangegeven onderverdeling naar ernstniveaus.

Tabel 5 Percentage mannen en vrouwen van 16 jaar en ouder naar score op de MHI-schaal

score	mannen	vrouwen
0-16 (ernstig ongezond)	0.3	0.9
17-36 (matig ongezond)	2.2	4.0
37-52 (mild ongezond)	6.4	10.1
53 – 100 (gezond)	91.1	85.1
Totaal	100 (N=971)	100 (N= 756)

* percentages afgrond op 1 decimaal

Indien we een verdeling toepassen conform de dichotomisering van Brønnum-Hansen en Rasmussen (Brønnum-Hansen & Rasmussen, 1996) (een kleiner of gelijk aan de waarde 52 wordt als ongezond beschouwd), dan blijkt dat voor mannen (vanaf 16 jaar) de (punt)prevalentie van mentale ongezondheid 8.9% is en voor vrouwen vanaf die leeftijd 14.9%.

In de bijlage worden de eerste resultaten gegeven van de berekening van een LGGG gebaseerd op bovenstaande afkappunten.

De SF-36 is ook afgenomen in het kader van het NEMESIS-project (Bijl et al., 1997b; Bijl et al., 1997a). De NEMESIS-studie (Netherlands Mental Health Survey and Incidence Study) is opgezet om kennis te verzamelen over de incidentie, prevalentie en het beloop van psychiatrische morbiditeit onder Nederlandse volwassenen (18-64 jaar), en de determinanten en de consequenties van psychische ziekten. Binnen de NEMESIS-studie is een volledig psychiatrisch interview afgenomen bij de respondenten. Hoewel binnen de NEMESIS-studie gegevens over 7076 personen bekend zijn, is het nadeel dat het alleen de leeftijdsgroep 18 tot en met 64 jaar betreft. Met name het ontbreken van gegevens over de leeftijdsgroep vanaf 65 jaar vormt voor de berekening van de gezonde levensverwachting een probleem. Berekening van de LGGG op basis van het NEMESIS-bestand kan dan in principe alleen op basis van een partiële overlevingstafel, hetgeen verwarring bij de interpretatie kan opleveren (Perenboom & van de Water, 1997a). Wellicht kan de prevalentie van ongezondheid voor 65 plussers geschat worden op basis van de gegevens over jongere leeftijdsgroepen.

Overigens leent deze NEMESIS studie zich voor een nadere bepaling van afkappunten en wegingsfactoren van de MHI-5. Hierop wordt in hoofdstuk 4 nader teruggekomen.

In de Morgenstudie (uitgevoerd door het RIVM in Doetinchem eo, Maastricht eo en Amsterdam) zijn met behulp van de SF-36 gegevens verzameld onder mannen en vrouwen van 20 tot 60 jaar. De prevalentie in deze groep van geestelijke ongezondheid staat weergegeven in tabel 6.

Tabel 6 Prevalentie geestelijke (on)gezondheid voor mannen en vrouwen, naar ernst van (on)gezondheid [onpubliceerde cijfers uit de MORGEN studie, zie ook Edlinger et al., 1998]

score	mannen	vrouwen
0-16 (ernstig ongezond)	0,4	0,7
17-36 (matig ongezond)	2,6	3,7
37-52 (mild ongezond)	7,4	11,6
53 – 100 (gezond)	89,5	84,0
totaal	100% (N= 6076)	100% (N=7438)

* percentages afgerond op 1 decimaal

In de MORGEN studie is de prevalentie van geestelijke ongezondheid iets hoger dan uit de TNO-studie naar voren komt (bij mannen 10,5% ongezondheid in de MORGEN-studie tegen 8,9% in de

TNO-studie, bij vrouwen respectievelijk 16% en 14,9%). Een duidelijke oorzaak van deze verschillen is niet aan te geven.

Gebruik van de gegevens van de MORGEN studie voor de berekening van een LGGG is, net als de gegevens uit de NEMESIS studie door de beperkte leeftijdsrange enigszins problematisch. In principe kan hiermee de LGGG uitsluitend berekend worden tot de leeftijd van 60 jaar. Anderszijds lijkt het mogelijk om de prevalentie van geestelijke ongezondheid voor de leeftijdsgroep boven de 60 jaar te schatten op basis van de gegevens uit de jongere leeftijdsgroepen.

Tenslotte zijn ook gegevens verzameld onder een steekproef van de Amsterdamse volwassen populatie, waarbij SF-36 gegevens verzameld zijn in een survey naar het gebruik van geoorloofde en ongeoorloofde drugs (Aaronson et al., 1998). Voor deze studie is een steekproef van 8686 personen van 12 jaar en ouder getrokken uit het gemeentelijk bevolkingsregister. De respons was 50.2% (4364 personen). 192 personen waren jonger dan 16 jaar. Omdat de nonrespons erg hoog was (hoewel niet hoger dan gebruikelijk in Amsterdam) is een apart nonrespons onderzoek gepleegd, waaruit bleek dat de nonrespons groep niet substantieel afweek van de respons groep op bijvoorbeeld de algemene gezondheid.

3.2 Gegevensbestanden m.b.t. de GHQ-12

De meest recente landelijk toepassing van de GHQ-12 heeft plaatsgevonden in het NEMESIS-onderzoek van het Trimbos-instituut (Bijl et al., 1997b; Bijl et al., 1997a) (zie paragraaf 3.1). De GHQ-12 is in deze studie toegepast om informatie te verkrijgen bij personen die weigerden aan het psychiatrische interview mee te doen. Zoals in paragraaf 3.1 reeds is aangegeven zijn de gegevens in de NEMESIS studie beperkt tot de leeftijdsgroep 18 tot en met 64 jaar. Gezien de relatie die er is tussen de GHQ-score en de leeftijd, is het niet zondermeer mogelijk om gegevens van jongere leeftijdsgroepen te gebruiken om de prevalentie in oudere leeftijdsgroepen te schatten.

Buiten deze NEMESIS studie is geen recent landelijk onderzoek bekend waarin de GHQ-12 is toegepast in Nederland. Wel is in de NIVEL Nationale studie naar Ziekten en Verrichtingen in de Huisartsen praktijk uit 1988 (Foets & Sixma, 1991) de GHQ-30 toegepast. Deze GHQ-30 omvat alle items die in de GHQ-12 zijn opgenomen, zoals in hoofdstuk 2.2 is beschreven. In de Nationale studie is geen boven of ondergrens gesteld aan de leeftijd, zodat deze gegevens zich lenen voor

berekening van een LGGG. Er zijn plannen om in 1999-2000 deze studie te herhalen, waarbij ook de GHQ-12 opgenomen wordt. Hiermee opent zich de mogelijkheid om op basis van de GHQ een complete LGGG te berekenen.

Verder wordt de GHQ-12 in diverse regionale onderzoeken door GGD'en toegepast (blijkens een mondelinge inventarisatie van lopende studies door een werkgroep van GGD-epidemiologen). Maar deze studies hebben als nadeel dat de dataverzameling plaats heeft gevonden bij specifieke groepen respondenten, waarbij veelal een bovenste leeftijdsgrens rond de 65 of 70 jaar is gesteld.

Voor een voorbeeld van een bestand m.b.t. de GHQ-12 zij verwezen naar de VTV 1997, deel III (Perenboom & van de Water, 1997a).

3.3 Gegevensbestanden m.b.t. de ABS

Sinds 1989 zijn de vijf 'negatieve' items (NAS) van de ABS opgenomen in de Continue Gezondheidsenquête van het CBS. De NAS is opgenomen in het schriftelijke deel van de gezondheidsenquête dat wordt afgenomen bij de niet-institutionele bevolking van 16 jaar en ouder. Voor de jaren 1989 tot en met 1995 is met behulp van de gegevens uit de gezondheidsenquêtes een LGGG berekend (Perenboom & van de Water, 1997a; Perenboom & van de Water, 1997b). Er zijn geen (recente) andere landelijke bestanden bekend waarin de ABS of onderdelen daarvan zijn opgenomen.

4 Conclusies en aanbevelingen

In VTV-1997 is voor het eerst een schatting gemaakt van de Levensverwachting in Goede Geestelijke Gezondheid in Nederland. Hierbij is gebruik gemaakt van de 5 negatieve items van de Affect Balance Scale (ABS), zoals opgenomen in de gezondheidsenquête van het CBS.

In bovenstaande rapportage is geïnventariseerd welke andere instrumenten in aanmerking komen als alternatief voor de ABS, gegeven het feit dat de ABS, hoewel veel toegepast, mede door de verschillende verschijningsvormen, niet is gevalideerd.

Uitgangspunt voor deze inventarisatie is geweest dat het moest gaan om gevalideerde, internationaal toegepaste en in het Nederlands vertaalde instrumenten. Op basis van deze criteria zijn uiteindelijk slechts twee instrumenten in aanmerking gekomen, namelijk de General Health Questionnaire (GHQ) en de Mental Health Inventory (5-item versie), zoals opgenomen in de SF-36 (MHI-5). In dit rapport zijn de kenmerken van de verschillende instrumenten (MHI-5, GHQ en de reeds toegepaste ABS) nog eens naast elkaar gezet. Bovendien is bekeken of er in Nederland bestanden zijn, waarvan gebruik gemaakt kan worden voor de berekening van een LGGG.

Op basis van inhoudelijke overwegingen lijkt toepassing van de MHI-5 het meest het overwegen waard. Zoals gezegd, de ABS heeft als nadeel dat het een onvoldoende onderbouwd instrument is. De GHQ heeft als nadeel dat het gevoelig is voor leeftijd. De MHI-5 (als onderdeel van de SF-36) kent deze nadelen niet, althans er zijn in de literatuur geen aanwijzingen voor gevonden.

Indien we kijken naar de beschikbaarheid van gegevens in Nederland, blijkt dat voor de ABS en voor de MHI-5 recente gegevensbestanden aanwezig zijn die betrekking hebben op de Nederlandse populatie vanaf circa 16 jaar. Voor wat betreft de GHQ-12 (en andere versies) zijn alleen gegevens beschikbaar voor specifieke leeftijdsgroepen (bijvoorbeeld voor de leeftijdsgroep 16 t/m 70 jaar). Wel zijn er bij het NIVEL plannen om in 1999-2000 de Nationale studie naar ziekten en verrichtingen in de huisartspraktijk te herhalen. In deze studie zal ook de GHQ-12 opgenomen worden.

Voordeel (tot op heden) van het gebruik van de ABS in de gezondheidsenquête is dat het mogelijk is om trends over de tijd weer te geven. Voor de MHI-5 is het op dit moment alleen mogelijk om een weergave voor 1996 te geven, gezien het feit dat de dataverzameling toen heeft plaats gevonden.

Indien evenwel de SF-36 een plaats zou krijgen in de POLS (Permanent Onderzoek Leefsituatie, waarin opgenomen is de gezondheidsenquête) van het CBS, lijkt de MHI-5 in de toekomst de beste keuze.

Om de mogelijkheden en de waarde van de MHI-5 voor de berekening van een genuanceerde en gewogen Levensverwachting in Goede Geestelijke Gezondheid te benutten is echter nader onderzoek noodzakelijk, waarbij de nadruk gelegd dient te worden op de bepaling van de juiste afkappunten voor de verschillende ernstniveaus en op de bepaling van de daarbij behorende gewichten. Voor dit nadere onderzoek leent zich een analyse op het NEMESIS-bestand, waarin zowel de SF-36 als de GHQ-12 als diagnostisch interviews (CIDI en SCID) zijn toegepast.

De concrete vragen en problemen die met behulp van een nadere analyse dienen te worden beantwoord respectievelijk opgelost zijn de volgende:

Is het met behulp van het NEMESIS-bestand mogelijk via een analyse van de MHI-5 en de GHQ-12 enerzijds en de CIDI en de SCID anderzijds om:

1. een afkappunt voor geestelijke ongezondheid te definiëren respectievelijk te valideren voor de MHI-5 en de GHQ-12?
2. de scoreranges van de MHI-5 respectievelijk de GHQ-12 die op basis van het afkappunt uit vraag 1 wijzen op ongezondheid onder te verdelen in drie ernstniveaus: milde, matige en ernstige geestelijke ongezondheid?
3. aan ieder ernstniveau een wegingsfactor toe te kennen, op basis waarvan een 'mental Health Adjusted Life Expectancy' (mHALE³) berekend kan worden?

ad 1. In de eerste plaats moet bekeken worden of in Nederland de subschaalscores van de MHI-5 van 52 en hoger inderdaad een normale tot goede geestelijke gezondheid representeren, zoals de ontwerpers van de MHI-5 aangeven en zoals Berwick et al. in hun studie aantreffen (Ware et al., 1997; Berwick et al., 1991). Het lijkt daarom zinvol om de analyse zoals door Berwick et al uitgevoerd (ROC-analyse) te herhalen voor de Nederlandse situatie. Daarnaast lijkt het verstandig om ook het afkappunt van de GHQ-12 nog eens opnieuw te bepalen.

ad 2. Binnen de scorerange van de MHI-5 worden min of meer arbitrair vier niveaus van (on)gezondheid gedefinieerd: 0-16 ernstig ongezond, 17-36 matig ongezond, 37-52 mild ongezond

³ geestelijk Gezond Levensjaar Equivalent (gGLE)

en 53-100: gezond. Door vergelijking van deze groepen met bijvoorbeeld de resultaten van de CIDI en SCID is de plausibiliteit van deze ernstniveaus te bekijken.

Bij de GHQ worden ook extra afkappunten gedefinieerd en bestudeerd in relatie tot de CIDI en SCID om te bezien of een verdere onderverdeling in ernstniveaus is aan te brengen.

ad 3. Om vervolgens aan de ernstniveaus binnen zowel de GHQ-12 als de MHI-5 wegingsfactoren toe te kennen zijn 3 opties aanwezig:

Optie A: Per ernstniveau wordt bepaald hoe de verdeling van specifieke (psychiatrische) diagnoses is. Met behulp van de bij die specifieke diagnoses behorende wegingsfactoren, kan een bij dat ernstniveau behorende 'totaal' wegingsfactor berekend worden.

Nadeel van deze optie is dat de generieke geestelijke gezondheid niet meer duidelijk onderscheiden is van psychiatrische aandoeningen.

Optie B: Een andere mogelijkheid is om per ernstniveau de verdeling van de ervaren (on)gezondheid te bepalen. De wegingsfactor voor ieder ernstniveau wordt dan berekend uit de reeds bekende wegingsfactoren voor ervaren gezondheid (uit VTV-1997).

Nadeel van deze optie is dat geestelijke gezondheid iets anders is dan ervaren gezondheid.

Optie C: De laatste optie is om per ernstniveau een specifieke omschrijving weer te geven (zoals voortdurend zenuwachtig, vaak in de put, zelden kalm en rustig, vaak neerslachtig en somber, en soms gelukkig) en dit door een panel te laten wegen.

Tenslotte zijn er nog enkele specifieke vragen die in een nadere analyse onderzocht kunnen worden of op een andere wijze opgelost moeten worden:

1. Is het mogelijk om in de scorereange boven waarde 52 van de MHI-5 - 'hetgeen zou wijzen op een minimaal normale gezondheid' - toch nog ernstniveaus te onderscheiden, bijvoorbeeld doordat ook in dat gebied een samenhang met een hogere kans op specifieke (psychiatrische) diagnoses aanwezig is?
2. NEMESIS heeft een beperkte leeftijdsrange (18 tot en met 64 jaar). Het is mogelijk dat bij ouderen de relatie tussen MHI en psychiatrische diagnoses afwijkt. Anderzijds is uit studies als van McCabe et al bekend dat er geen relatie aanwezig lijkt tussen de MHI-5 en de leeftijd.

Met bovenstaande nadere studies kan Nederland een trend zetten op het terrein van de LGGG en de mogelijkheid openen om op diverse manieren een vergelijking aan te gaan met LGGG's die in het buitenland berekend worden.

Literatuur

AARONSON NK, ACQUADRO C, ALONSO J, et al. International quality of life assessment (IQOLA) project. *Quality of Life Research* 1992; 1:349-51.

AARONSON NK, MULLER M, COHEN PDA, et al. Translation, validation and norming of the Dutch Language version of the SF-36 Health Survey in community and chronic disease populations. concept : Amsterdam/Leiden.

BERWICK DM, MURPHY JM, GOLDMAN PA, et al. Performance of a five-item mental health screening test. *Medical Care* 1991; 29:169-76.

BIJL RV, ZESSEN G van, RAVELLI A. Psychiatrische morbiditeit onder volwassenen in Nederland: het NEMESIS-onderzoek. II. Prevalentie van psychiatrische stoornissen. *Nederlands Tijdschrift voor Geneeskunde* 1997a; 141:(50):2453-60.

BIJL RV, ZESSEN G van, RAVELLI A, et al. Psychiatrische morbiditeit onder volwassenen in Nederland: het NEMESIS-onderzoek. I. Doelstelling, opzet en methoden. *Nederlands Tijdschrift voor Geneeskunde* 1997b; 141:(50):2448-52.

BRADBURN NM. The structure of psychological well-being. Chicago: Aldine Publishing, 1969.

BRADBURN NM, CAPLOVITZ D. Affect Balance Scale (ABS). In: Furer JW, König-Zahn C, Tax B (eds.) *Beschrijving en evaluatie van vragenlijsten: 3: Psychische gezondheid*. Assen: Van Gorcum, 1995: pp.42-54.

BRAZIER JE, WALTERS SJ, NICHOLL JP, et al. Using the SF-36 and Euroqol on an elderly population. *Quality of Life Research* 1995; 5:195-204.

BRØNNUM-HANSEN H, RASMUSSEN NKr. Mental health expectancy in Denmark 1994. Paper presented at the 2-nd annual meeting of Euro-REVES, London, 4-5 july, 1996. DICE: Copenhagen, 1996:

EDLINGER M, HOEYMANS N, TIJHUIS M, ET AL. De kwaliteit van leven (RANS-36) in twee Nederlandse populaties. *Tijdschrift voor Gezondheidswetenschappen* 1998; 76: 211-9.

FINLAY-JONES RA, MURPHEY E. Severity of psychiatric disorder and the 30-item General Health Questionnaire. *Brittish Journal of Psychiatry* 1979; 134:609-16.

GISPERT R, RITCHIE K, RAJMIL L, RUÉ M, et al. Mental health expectancy: an indicator to bridge the gap between psychiatric and public health perspectives about the population's health. *Acta Psychiatrica Scandinavica* 1998;96 (in press).

GISPERT R, RAJMIL LA, RUÉ M, et al. Mental Health Expectancy in Catalonia (Spain), 1994. Paper presented at the 2-nd EuroREVES meeting, London England, July 1996:

GOLDBERG DP. General Health Questionnaire (GHQ). In: Furer JW, König-Zahn C, Tax B (eds.) Beschrijving en evaluatie van vragenlijsten: 3: Psychische gezondheid. Assen: Van Gorcum, 1995: pp.104-18.

GOLDBERG DP, WILLIAMS P. A user's guide to the general health questionnaire. 2. Windsor, Berkshire: The NFER-NELSON Publishing Company, 1988.

GOODCHILD ME, DUNCAN-JONES P. Chronicity and the General Health Questionnaire. British Journal of Psychiatry 1985; 146:55-61.

HEYDENDAEL PHJM, FURER JW, HODIAMONT PPG, et al. De waarde van vijf screeningstests voor geestelijke (on)gezondheid. Gezondheid & Samenleving 1986; 7:(3):138-44.

HIMMELFARB S, MURRELL SA. Reliability and validity of five mental health scales in older persons. J Gerontol 1983 May 1998; 38:333-9.

JENKINSON C, LAYTE R. Development and testing of the UK SF-12. Journal of Health Services Research and Policy 1997; 2:(1):14-8.

JOHNSON JA, COONS SJ. Comparison of the EQ-5D and SF-12 in an adult US sample. Quality of Life Research 1998; 7:155-66.

KEMPEN GIJM, ORMEL J. [Measuring psychological well-being in the elderly]. Tijdschr Gerontol Geriatr 1992 Dec 1998; 23:225-35.

KOETER MWJ, ORMEL J. General Health Questionnaire: Nederlandse bewerking/Handleiding. 1991:

MCCABE CJ, THOMAS KJ, BRAZIER JE, et al. Measuring the mental health status of a population: a comparison of the GHQ-12 and the SF-36 (MHI-5). Br J Psychiatry 1996; 169 :516-21.

MUTAFOVA M, WATER HPA van de, MALESHKOV C, et al. Attempt for the assessment of the mental health of the population in Bulgaria. Paper, presented at REVES 9, Rome, December 11-13, 1996.

MUTAFOVA M, WATER HPA van de, MALESHKOV C, et al. Attempt for the assessment of the mental health of the population in Bulgaria. REVES 9 Proceedings. Rome: ISTAT, 1998:

PERENBOOM RJM, HERTEN LM van, BOSHUIZEN HC, et al. Trends in de gezonde levensverwachting in Nederland 1983-1994, met een verdeling naar ernst van ongezondheid. In:

Maas PJ van der, Kramers PGN (eds.) Volksgezondheid Toekomst Verkenning 1997; deel III: Gezondheid en Levensverwachting gewogen. Maarssen: Elsevier/de Tijdstroom, 1997a: pp.53-77.

PERENBOOM RJM, HERTEN LM van, BOSHUIZEN HC, et al. Trends in de Gezonde Levensverwachting in Nederland, 1983-1994, met een verdeling naar ernst van ongezondheid. Leiden: TNO Preventie en Gezondheid, 1997b.

PERENBOOM RJM, WATER HPA van de. Levensverwachting in goede geestelijke gezondheid in Nederland, 1989-1995: een eerste proeve. In: Maas PJ van der, Kramers PGN (eds.) Volksgezondheid Toekomst Verkenning 1997: deel III: Gezondheid en Levensverwachting gewogen. Maarssen: Elsevier/de Tijdstroom, 1997a: pp.139-52.

PERENBOOM RJM, WATER HPA van de. Levensverwachting in goede geestelijke gezondheid in Nederland, 1989-1995: een eerste proeve. Leiden: TNO Preventie en Gezondheid, 1997b.

STACEY CA, GATZ M. Cross-sectional age differences and longitudinal change on the Bradburn Affect Balance Scale. *Journals of Gerontology* 1991; 46:(2):P76-P78

STEWART AL, WARE JE. De Medical Outcomes Study Short-form General Health Survey (MOS SF-20). In: Furer JW, König-Zahn C, Tax B (eds.) Het meten van de Gezondheidstoestand: beschrijving en evaluatie van vragenlijsten. Deel 1 Algemene Gezondheid. Assen: Van Gorcum, 1993: pp.77-85.

STOOTHARD MEA, ESSINK-BOT ML, BONSEL GJ et al. Wegingsfactoren voor ziekte. In: Maas PJ van der, Kramers PGN (eds.) Volksgezondheid Toekomst Verkenning 1997: deel III: Gezondheid en Levensverwachting gewogen. Maarssen: Elsevier/de Tijdstroom, 1997a: pp.153-78.

VAN DER DOES FE, DE NEELING JN, SNOEK FJ, et al. Symptoms and well-being in relation to glycemic control in type II diabetes. *Diabetes Care* 1996 Mar 1998; 19:204-10.

WARE JE, KOSINSKI M, BAYLISS MS, et al. Comparison of methods for the scoring and statistical analysis of SF-36 health profile and summary measures: summary of results from the Medical Outcomes Study. *Medical Care* 1995; 33:AS264-AS279

WARE JE, KOSINSKI M, KELLER SD. A 12-item Short-form health survey: construction of scales and preliminary test of reliability and validity. *Medical Care* 1996; 34:220-6.

WARE JE, SHERBOURNE CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992; 30:473-83.

WARE JE, SNOW KK, KOSINSKI M, et al. SF-36 Health Survey: Manual and interpretation guide. 2. Boston: The Health Institute, New England Medical Center, 1997.

ZEE KI van der, SANDERMAN R. Het meten van de algemene gezondheidstoestand met de RAND-36: een handleiding. Noordelijk centrum voor Gezondheidsvraagstukken, RUG: Groningen, 1993.

Bijlage

In deze bijlage worden eerste resultaten weergegeven van een LGGG, gebaseerd op de MHI-5. Hiervoor is gebruik gemaakt van de gegevens uit het TNO-gegevensbestand van de SF-36 (zie hoofdstuk 3.1). Voor de berekening van de LGGG is de zogenaamde methode van Sullivan toegepast, waarbij de totale levensverwachting met behulp van prevalentiegegevens verdeeld wordt in gezonde en ongezonde jaren.

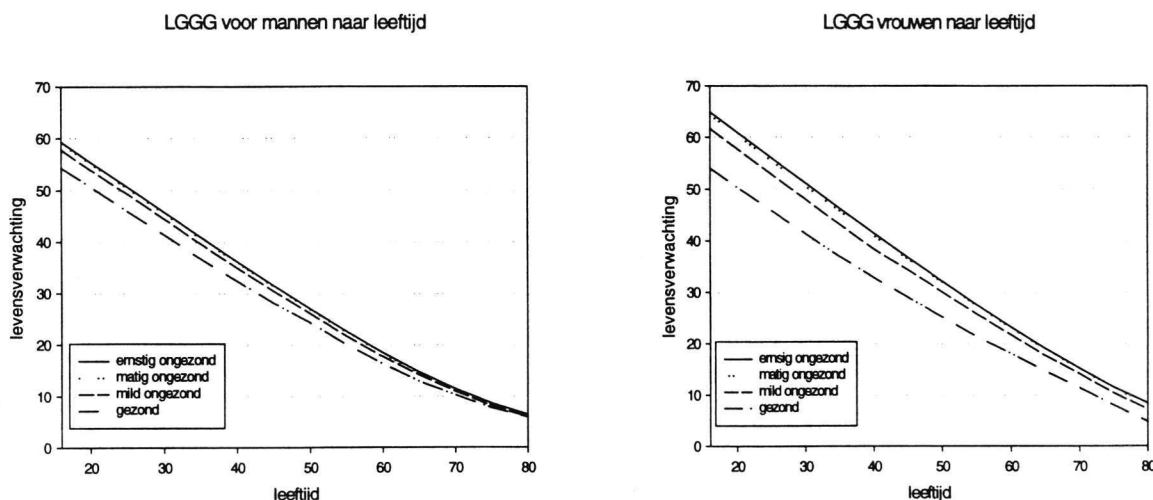
Om de totale levensverwachting te berekenen zijn (bij nog ontbreken van gegevens over 1996) sterftecijfers van 1995 gebruikt. Daarnaast is verondersteld dat de prevalentie van geestelijke (on)gezondheid in de intramurale populatie gelijk is aan die in de zelfstandig wonende populatie, dus gelijk aan de cijfers die uit de TNO-studie naar voren komen.

In tabel b1 is de LGGG voor mannen en vrouwen weergegeven voor de leeftijd van 16 jaar en voor de leeftijd van 65 jaar. In figuur b1 worden de resultaten voor alle leeftijden grafisch weergegeven.

Tabel b1 Totale levensverwachting en te verwachten jaren in ongezondheid en gezondheid voor mannen en vrouwen van 16 en 65 jaar, 1995

leeftijd	totale levensverwachting	jaren met milde ongezondheid	jaren met matige ongezondheid	jaren met ernstige ongezondheid	gezonde jaren
<i>mannen</i>					
16 jaar	59,3	3,5	1,2	0,3	54,4
65 jaar	14,7	1,2	0,4	0,2	12,9
<i>vrouwen</i>					
16 jaar	64,9	7,5	2,7	0,6	54,1
65 jaar	19,1	3,0	0,9	0,1	11,4

N.B. Cijfers afgerond op 1 decimaal



Figuur 1 Levensverwachting en jaren in gezondheid en ongezondheid voor mannen (links) en vrouwen (rechts), voor verschillende leeftijden, 1995

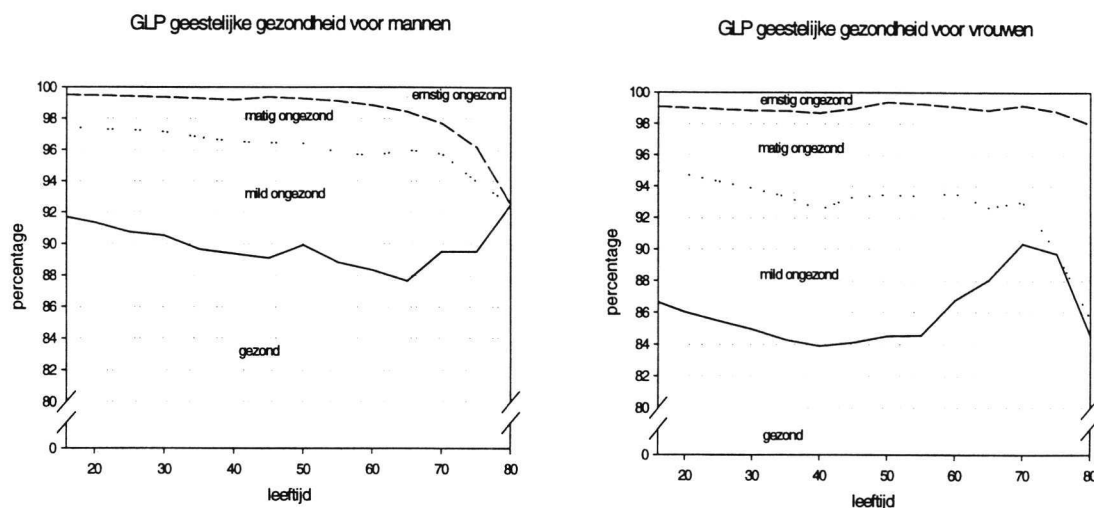
In figuur 1 wordt links voor mannen en rechts voor vrouwen voor verschillende leeftijden de Levensverwachting in goede geestelijke gezondheid en de te verwachten jaren in geestelijke ongezondheid weergegeven. De bovenste curve op beide grafieken geeft de totale levensverwachting weer. De ruimte tussen de bovenste en de op een na bovenste curve geeft het aantal jaren in ernstige ongezondheid weer. De ruimte tussen de op een na bovenste en de op twee na bovenste curve geeft de jaren in matige ongezondheid weer, terwijl die tussen de op twee na bovenste en de onderste curve de jaren in milde ongezondheid weer geven. De onderste curve geeft het aantal jaren in goede gezondheid (de levensverwachting in goede geestelijke gezondheid, LGGG) weer.

Op 16 jarige leeftijd is de totale levensverwachting voor mannen 59,3 jaar en voor vrouwen 64,9 jaar (in 1995). Voor mannen is de LGGG op 16 jarige leeftijd 54,4 jaar, voor vrouwen 54,1 jaar. Mannen van 16 jaar kunnen verwachten circa 3,5 jaar in milde ongezondheid door te brengen, voor vrouwen is dat ruim 7,5 jaar. De te verwachten periode in matige ongezondheid is voor mannen van 16 jaar 1,2 jaar, voor vrouwen 2,7 jaar, terwijl het aantal te verwachten jaren in ernstige ongezondheid voor mannen van 16 jaar 0,3 jaar is (de curven lijken ook nagenoeg samen te vallen) en voor vrouwen 0,6 jaar.

Hoewel vrouwen die in 1995 16 jaar waren dus konden verwachten ruim 5,5 jaar langer te leven dan mannen, hebben vrouwen zelfs wat minder jaren in goede geestelijke gezondheid te verwachten dan mannen. De voorsprong die vrouwen ten opzichte van mannen hebben in totale levensverwachting wordt dus volledig te niet gedaan, wanneer de gezondheidsstatus in de vergelijking wordt betrokken.

Dit blijkt ook wanneer we de LGGG presenteren als proportie of percentage van de totale levensverwachting (Gezond LevensPercentage voor de Geestelijke Gezondheid GLPGG).

Figuur 2 Jaren in gezondheid en ongezondheid als proportie van de totale levensverwachting voor mannen (links) en vrouwen (rechts) voor verschillende leeftijden (1995)



Voor mannen van 16 jaar is de GLPGG 91,8%, terwijl de GLPGG voor vrouwen 86,6% bedraagt. Mannen kunnen verwachten ruim 8% van hun leven in geestelijke ongezondheid door te brengen, vrouwen 13,5%. Dit percentage komt overigens sterk overeen met de ruwe prevalentie van geestelijke ongezondheid, zoals beschreven in hoofdstuk 3.1. Deze gelijkenis wordt veroorzaakt doordat de prevalentie van geestelijke ongezondheid, gemeten met de MHI-5, voor iedere leeftijd nagenoeg gelijk is. Indien we de GLPGG bekijken voor mannen en vrouwen van bijvoorbeeld 80 jarige leeftijd, zien we dat deze voor mannen circa 92% is en voor vrouwen circa 84%. Er is voor de verschillende leeftijden wel enige variatie te onderkennen in GLPGG, maar er is geen afnemend patroon zoals bij andere gezondheidsconcepten aanwezig. De fluctuerende curven die het patroon van het GLPGG op hogere leeftijd weergeven, wordt met name veroorzaakt door een enigszins fluctuerende prevalentie, die weer mede veroorzaakt wordt door de kleine N in de hogere leeftijdsgroepen.

Opvallend is dat ook bij de geestelijke gezondheid de voorsprong van vrouwen ten opzichte van mannen in levensverwachting door de aanwezigheid van meer ongezonde jaren in het niet verdwijnt.

Dit fenomeen treffen we ook aan als we de GLV op basis van aanwezige beperkingen beschouwen. Indien de GLV op basis van ervaren gezondheid wordt berekend (voor beide geslachten circa 60 jaar) blijkt het aantal jaren in goede gezondheid voor beide geslachten nagenoeg gelijk te zijn (Perenboom & van de Water, 1997b).

Al met al kan hieruit geconcludeerd worden dat de extra jaren die vrouwen kunnen verwachten te leven, nagenoeg geheel in een staat van (meestal milde) ongezondheid worden doorgebracht.