

# Een multicenterstudie naar betrouwbaarheid, validiteit en gevoeligheid voor verandering van de HONOS65+ binnen de ouderenpsychiatrie

M. J. AARTSEN, A. K. SPITSBAARD, C. VAN BAARSEN, A. D. F. DHONDT, M. MASCINI, A. NEFS, A. SNOEIJERS, M. VAN DIEST, R. WILTING

**ACHTERGROND** In de geestelijke gezondheidszorg voor ouderen is er een groeiende behoefte om resultaten van klinische behandeling inzichtelijk te maken en te evalueren. De Health of the Nations Outcome Scales 65+ (HONOS65+) lijkt daarvoor een geschikt instrument, maar de validiteit ervan bij Nederlandse ouderen is nog onbekend.

**DOEL** Vaststellen van betrouwbaarheid, validiteit en gevoeligheid voor verandering van de HONOS65+ bij ouderen met psychiatrische aandoeningen.

**METHODE** Het biopsychosociaal functioneren van ouderen van 60 jaar en ouder ( $n = 168$ ) werd vastgesteld aan de hand van bestaande en gevalideerde meetinstrumenten, en afgezet tegen resultaten met de HONOS65+. Na drie maanden werd bij dezelfde steekproef de mate van verandering vastgesteld.

**RESULTATEN** Er werd een voldoende tot goede interbeoordelaarsbetrouwbaarheid gevonden voor 5 van de 9 beoordeelde subschalen van de HONOS65+, te weten opzettelijke verwonding, alcoholmisbruik, somatische problemen, problemen met depressieve symptomen en problemen met algemene dagelijkse levensverrichtingen. Op grond van de gegevens van 116 cliënten die na drie maanden opnieuw onderzocht werden, bleek de HONOS65+ gevoelig voor verandering.

**CONCLUSIE** De HONOS65+ is een betrouwbaar en valide instrument bij ouderen met affectieve stoornissen en kan verandering van problemen vaststellen. Conclusies over de meeteigenschappen van de HONOS65+ bij ouderen met andere psychiatrische aandoeningen konden niet worden getrokken vanwege te kleine aantallen.

[TIJDSCHRIFT VOOR PSYCHIATRIE 52(2010)8, 543-553]

**TREFWOORDEN** HONOS65+, ouderenpsychiatrie, validatieonderzoek

Bij instellingen voor geestelijke gezondheidszorg is er een toenemende behoefte om de resultaten van klinische behandelingen inzichtelijk te maken en te evalueren. Inzichten in de effecten van behandelingen geven niet alleen de mogelijkheid om kosteneffectiviteit van behandelingen in kaart te brengen, maar ook om ineffektieve behandelingen tijdig op te sporen en

waar mogelijk bij te stellen. Er bestaan verschillende meetinstrumenten die daarvoor benut kunnen worden, waaronder de Health of the Nations Outcome Scales (HONOS). De HONOS is een samengestelde schaal die op een eenvoudige wijze klinische uitkomstmaten zoals mentale en fysieke gezondheid en sociaal functioneren in kaart brengt (Wing e.a. 1998), en gevalideerd is voor de

volwassen psychiatrische populatie (Mulder e.a. 2004).

Deze schaal is vanwege het frequent voorkomen van fysieke en cognitieve problemen echter minder geschikt voor mensen van 65 jaar of ouder (Shergill e.a. 1999). Om die reden is er een aangepaste HONOS ontwikkeld, de HONOS65+, waarbij er aandacht is voor specifieke problemen bij ouderen zoals mate van bewustzijn, een specificatie van aspecten van depressie zoals schuldgevoelens en spelen met gedachten over zelfdoding en een specificatie van de chroniciteit van cognitieve problemen (zie voor een volledig overzicht van aanpassingen Burns e.a. 1999). Inmiddels is ook de 65+-versie van de HONOS in het Nederlands vertaald (Staring e.a. 2003).

Uit Gronings onderzoek is gebleken dat hulpverleners de HONOS65+ ervaren als een bruikbaar instrument, zowel op patiënt- als op groepsniveau, en dat het de problemen op relevante levensgebieden binnen de ouderenpsychiatrie duidelijk maakt (Broersma & Sytema 2008). Broersma en Sytema hebben in hun onderzoek het toetsen van de betrouwbaarheid, validiteit en gevoeligheid voor verandering buiten beschouwing gelaten, zodat zij daarover geen uitspraken konden doen.

Het doel van onze studie is derhalve het vaststellen van de betrouwbaarheid, validiteit, en gevoeligheid voor veranderingen van de HONOS65+ bij ouderen die ambulante of intramuraal behandeld worden door instellingen voor de geestelijke gezondheidszorg (ggz).

## METHODE

### Steekproef

De deelnemers aan de studie werden willekeurig geselecteerd uit het cliëntenbestand van 4 ggz-instellingen in Nederland. Omdat de HONOS65+ geschikt verondersteld wordt voor de totale populatie ouderen, kwamen in principe alle cliënten in de leeftijd van 65 jaar of ouder in aanmerking voor de studie. Aangezien verschillende

instellingen voor de afdeling Ouderenpsychiatrie een ondergrens van 60 jaar hanteren, sloten wij, omwille van de representativiteit van de onderzoeksresultaten voor de Nederlandse situatie, ouderen van 60 tot 65 jaar niet op voorhand uit.

Een aantal schalen die gebruikt worden om de concurrente validiteit vast te stellen zijn ongeschikt voor mensen met cognitieve problemen. Daarom hanteerden wij een score van 19 of hoger op de Mini Mental State Examination (MMSE; Folstein e.a. 1975) als ondergrens voor deelname (Burke e.a. 1989). Verder werd ook het beheersen en het begrijpen van de Nederlandse taal als voorwaarde gesteld en werden cliënten die volgens de behandelaren te kwetsbaar waren, niet benaderd voor deelname aan het onderzoek.

### Opzet

Cliënten die voldeden aan de inclusiecriteria werd gevraagd om deel te nemen aan de studie nadat hun het doel en de inhoud van de studie mondeling en schriftelijk waren uitgelegd. Toestemming voor deelname aan het onderzoek werd vastgelegd in een schriftelijke *informed consent*. De dataverzameling vond plaats tussen juni 2006 en december 2008. Het onderzoek werd goedgekeurd door de medisch-ethische commissie (GGZ Eindhoven) en de commissie wetenschap (Altrecht GGZ). Er werd twee keer gemeten met een tussenperiode van gemiddeld 3 maanden. Voor beide metingen ( $T_0$  en  $T_1$ ) werd de cliënt thuis of in de kliniek geïnterviewd door getrainde interviewers. In twee van de vier instellingen waren de getrainde interviewers tevens behandelaren. In de twee andere instellingen waren de interviewers niet betrokken bij de behandeling zodat de voorkennis van de cliënten beperkt was.

Voor het vaststellen van de interbeoordelaarsbetrouwbaarheid namen twee beoordelaars binnen een week en onafhankelijk van elkaar de HONOS65+ af bij dezelfde cliënten. Op basis van criteria van Shrout (1998) werd een intraklassecorrelatiecoëfficiënt (ICC) lager dan 0,40 als onvoldoende betrouwbaar beschouwd; een ICC van

0,41-0,60 als matig betrouwbaar; van 0,61-0,80 als voldoende betrouwbaar en boven de 0,81 als uitstekend betrouwbaar beschouwd.

Drie maanden na de uitgangsmeting werden de nog in behandeling zijnde cliënten opnieuw bezocht om een indruk te krijgen van de mate waarin de HONOS65+ gevoelig is voor verandering. Veranderingen in de totaalscore van de HONOS65+ werden vergeleken met veranderingen op de totaalscore van de *Brief Psychiatric Rating Scale* (BPRS) en het klinisch oordeel van de behandelaar.

### Bepalen van de validiteit

In de literatuur worden verschillende typen validiteit beschreven, maar de onderlinge verschillen zijn soms klein en naamgeving is niet altijd consistent. In dit onderzoek hanteerden we twee typen validiteit, concurrente validiteit en constructvaliditeit, gebaseerd op een klassiek en veel geciteerd werk van Cronbach en Meehl (1955). Onder concurrente validiteit verstonen wij de samenhang tussen de verschillende subschalen en andere gevalideerde instrumenten. Onder constructvaliditeit verstonen wij de samenhang tussen de gestelde diagnose en de hoogte van de score op de subschalen van de HONOS65+. De subschalen met betrekking tot psychiatrische klachten (1, 2, 6 en 8) van de HONOS65+ werden gevalideerd met de *Brief Psychiatric Rating Scale* (Overall & Gorham 1962); problemen met depressieve symptomen (subschaal 7) met de *Geriatric Depression Scale* (GDS; Yesavage e.a. 1983); cognitieve problemen (subschaal 4) werden gevalideerd met de *MMSE* (Folstein e.a. 1975); problemen met alcohol (subschaal 3) werden gevalideerd met de *Alcohol Use Disorder Test* (AUDIT; Saunders e.a. 1993); en problemen met activiteiten van het algemeen dagelijks leven (ADL-activiteiten) (10) werden gevalideerd met de *Barthel-ADL-index* (Mahony & Barthel 1965). Voor de subschalen 9 (problemen met sociale relaties), 11 (problemen met wonen) en 12 (problemen met dagactiviteiten) waren er geen valide subschalen beschikbaar waartegen ze betekenisvol konden worden afgezet. Bovendien heb-

ben deze subschalen in de intramurale setting een heel andere betekenis dan in de ambulante setting. Om die reden werden deze bij de beoordeling van de validiteit buiten beschouwing gelaten.

### Powerberekening

De minimale omvang van de steekproef werd gebaseerd op het gewenste significantieniveau ( $\alpha = 0,05$ ), de gewenste power ( $p = 0,80$ ) en de te verwachten correlaties met andere meetschalen. Omdat correlaties met andere schalen binnen deze doelgroep nog niet bekend waren, gebruikten wij bevindingen uit ander onderzoek als waarschijnlijkste schatters. In de studie van Burns e.a. (1999) onder ruim 200 patiënten die zowel ambulante als intramuraal werden behandeld, varieerden correlaties met de HONOS65+ van -0,23 (MMSE) tot 0,76 (BPRS). Wanneer we uitgingen van een minimale correlatie van 0,25 (een conservatieve schatting), was er een steekproef van minimaal 122 deelnemers nodig voor het toetsen van de validiteit. In overeenstemming met bevindingen van Burns e.a. (1999), waarin een uitstekende interbeoordelaarsbetrouwbaarheid gevonden werd (kappa 0,70-0,82), werden voor het toetsen van de betrouwbaarheid minimaal 30 cliënten door twee verschillende beoordelaars bezocht. Op basis van diezelfde studie waren voor het vaststellen van ten minste 30% verandering op de items bij een standaarddeviatie van 0,97 en een gemiddelde van 1,12 er ten minste 114 vervolgmetingen nodig. Rekening houdend met een uitval van 20% in 3 maanden streefden wij naar een steekproef van minimaal 150 cliënten. De interbeoordelaarsbetrouwbaarheid is bepaald aan de hand van de ICC.

### Meetinstrumenten

De HONOS65+ (Burns e.a. 1999) is een voor ouderen aangepaste versie van de HONOS (Wing e.a. 1998) en brengt de mate van problemen op diverse levensgebieden in kaart. De schaal bestaat uit 12 subschalen. Elke subschaal wordt gemeten op een vijfpuntsschaal die loopt van 0 (geen pro-

blemen) tot 4 (ernstige tot zeer ernstige problemen). De somscore van de HONOS65+ loopt van 0 (geen problemen op 12 levensgebieden) tot 48 (zeer veel problemen op alle 12 levensgebieden). De HONOS65+ is gebaseerd op zelfrapportage door de cliënt, observatie door de behandelaar en observatie van nauw betrokkenen.

De BPRS bestaat uit 16 items met de categorieën 0 (symptoom niet aanwezig) tot 7 (extreem ernstige stoornis). De totaalscore van de BPRS geeft een optelling van een breed scala aan psychiatrische stoornissen, zoals depressieve gevoelens, agitatie, angsten over lichamelijke gezondheid en hallucinaties. De somscore loopt van 16 (geen psychiatrische stoornis) tot 112 (zeer ernstige stoornis), en is deels gebaseerd op zelfrapportage en deels op observatie.

Problemen met alcoholgebruik werden in kaart gebracht met de AUDIT. De AUDIT is een zelfrapportageschaal bestaande uit 10 vragen met 5 antwoordcategorieën. De score loopt van 0 (geen alcoholisme) tot 41 (zeer waarschijnlijk alcoholisme). Een score van 8 of hoger suggereert problemen met alcoholgebruik.

De MMSE geeft een globale indruk van het cognitief functioneren en wordt onder andere gebruikt als screeningsinstrument voor dementie. De score op de MMSE loopt van 0 tot 30, een hogere score komt overeen met betere cognitieve functies. Een score tussen 18 en 24 wordt in verband gebracht met lichte cognitieve problemen, een score lager dan 18 met ernstige cognitieve problemen zoals dementie.

De GDS is bestemd voor het meten van depressieve symptomen bij ouderen. De schaal bestaat uit 30 dichotome (ja/nee) vragen zoals 'bent u innerlijk tevreden met uw leven?' en 'bent u bang dat u iets naars zal overkomen?'. De totaalscore loopt van 0 (geen symptomen) tot 30 (zeer ernstige depressieve symptomen). Een score van 11 of hoger is indicatief voor depressie. In dit onderzoek werden de vragen niet zelf ingevuld, maar voorgelezen en door de interviewer gescoord.

De Barthel-ADL-index brengt de behoefte aan fysieke ondersteuning in kaart, en loopt van 0

(op alle ADL-gebieden hulp nodig) tot 100 (op geen van de ADL-gebieden hulp nodig). De globale verandering ten opzichte van het begin van de behandeling volgens het oordeel van de behandelaar werd vastgesteld met vraag 2 van de Clinical Global Impression (CGI)-schaal (Guy 1976). Dit is een 7-puntsschaal die loopt van 1 (zeer veel verbeterd) tot 7 (zeer veel verslechterd).

## RESULTATEN

### Interview en steekproef

De dataverzameling vond plaats tussen juni 2006 en januari 2008. In totaal werden 168 cliënten geïnterviewd bij de beginmeting. De cliënten waren afkomstig uit de provincie Utrecht (Altrecht GGZ;  $n = 84$ ), de provincie Noord-Brabant (GGZ Eindhoven;  $n = 34$ ; en GGZ Westelijk Noord-Brabant;  $n = 30$ ) en de provincie Noord-Holland (GGZ Noord-Holland-Noord;  $n = 25$ ). Het basisinterview duurde gemiddeld 45 min (SD: 12,2, uitersten 20-120). Het interview tijdens follow-up duurde 29 min (SD: 11,2; uitersten 7,5-45). Verschillen in afnameduur tussen instellingen waren niet significant. Na ongeveer 3 maanden (mediaan: 3,3; SD: 1,3; uitersten: 2-11) konden 116 cliënten opnieuw ondervraagd worden. Uitval ( $n = 52$ ) werd veroorzaakt door verandering van behandelaar (42%), niet meer in behandeling zijn (34%, van wie de helft zodanig hersteld was dat behandeling niet verder nodig was), weigering (10%), overlijden (6%), te ernstige ziekte (4%), of een andere oorzaak (4%).

De gemiddelde leeftijd van de cliënten was 71,4 (SD: 7,10; 60-93) en 73% van de ondervraagden was vrouw. Het merendeel (67%) van de cliënten had een affectieve stoornis (angst, stemmingsstoornissen of combinaties daarvan) (zie verder tabel 1).

### Psychometrische eigenschappen

**Betrouwbaarheid** De resultaten van de interbeoordelaarsbetrouwbaarheid staan weer-

TABEL 1		Kenmerken van deelnemers aan onderzoek naar de HONOS65+ (n = 168)
Gemiddelde leeftijd in jaren; (SD); uitersten		71,4 (7,1); 60-93
Geslacht (% man)		26,8
Behandelsetting (% ambulante)		69,0
Opleidingsniveau		
Basisschool		33,5
Middelbare opleiding		45,5
Hogere opleiding		21,0
Burgerlijke staat		
Gehuwd		38,1
Verweduwd		30,4
Gescheiden		19,6
Nooit gehuwd geweest		10,7
Hoofddiagnose		
Affectieve stoornissen (angst- en stemmingsstoornissen) (%)		66,7
Middelenmisbruik (%)		2,4
Hallucinaties/wanen (%)		8,9
Somatoforme stoornis (%)		7,1
Anders (%)		5,4
Onbekend (%)		9,5

SD = standaarddeviatie

gegeven in tabel 2. De betrouwbaarheid van de verschillende items vertoonde grote variaties. Er werd een voldoende tot goede ( $ICC > 0,60$ ) interbeoordelaarsbetrouwbaarheid gevonden voor 5 van de 9 beoordeelde subschalen, te weten opzettelijke verwonding, alcoholmisbruik, somatische problemen, problemen met depressieve symptomen en problemen met ADL. De interbeoordelaarsbetrouwbaarheid was redelijk voor gedragsstoornissen ( $ICC = 0,46$ ), maar was laag en niet significant voor cognitieve problemen ( $ICC = 0,31$ ), hallucinaties ( $ICC = 0,19$ ) en andere psychische ziekten ( $ICC = 0,22$ ).

**Concurrente validiteit** In tabel 3 staan de resultaten voor de concurrente validiteit weergegeven. Gedragsstoornissen zoals gemeten met de HONOS65+ hingen samen met de score op de BPRS ( $r = 0,43$ ) en in mindere mate met de GDS ( $r = 0,31$ ), maar niet met die op de MMSE, de AUDIT en de Barthel-index. Opzettelijke zelfverwonding hing samen met de score op de GDS en de BPRS, maar niet met die op de MMSE en de AUDIT. Problemen met alcoholgebruik lieten een sterke samenhang zien met de score op de AUDIT ( $r = 0,50$ ), maar ook met die op de BPRS ( $r = 0,45$ ) en in mindere mate met die op de MMSE ( $r = 0,19$ ).

TABEL 2			Intraclasscorrelatiecoëfficiënt* (ICC) van de subschalen van de HONOS65+ (n = 32)
	Subschaal		ICC (95%-betrouwbaarheidsinterval)
1	Gedragsstoornissen		0,42 (0,10-0,67)
2	Opzettelijke zelfverwonding		0,63 (0,36-0,80)
3	Problemen met alcohol of drugsgebruik		0,93 (0,86-0,96)
4	Cognitieve problemen		0,32 (-0,03-0,59)
5	Problemen met somatische ziekten/handicap		0,62 (0,36-0,79)
6	Problemen die gepaard gaan met hallucinaties of wanen		0,11 (-0,24-0,44)
7	Problemen met depressieve symptomen		0,65 (0,40-0,81)
8	Andere geestelijke en gedragsproblemen		0,23 (-0,12-0,53)
10	Problemen met ADL-activiteiten		0,82 (0,66-0,91)

\*One-way random effect model.  
HONOS65+ = Health of the Nations Outcome Scales 65+; ADL = algemene dagelijkse levensverrichtingen.

Cognitieve problemen hingen alleen samen met de score op de MMSE ( $r = -0,32$ ). Somatische problemen hingen het sterkst samen met de score op de Barthel-index ( $r = -0,42$ ), maar ook met scores op de MMSE, de GDS en de BPRS. Problemen die gepaard gaan met hallucinaties of wanen hingen samen met de score op de BPRS ( $r = 0,20$ ) en die op de Barthel-index ( $r = -0,22$ ). Problemen met depressieve klachten hingen zeer sterk samen met de score op de GDS ( $r = 0,78$ ) en die op de BPRS ( $r = 0,61$ ). Andere geestelijke problemen hingen vooral samen met de score op de BPRS ( $r = 0,39$ ), maar ook met die op de GDS ( $r = 0,28$ ). Tot slot hingen problemen met ADL sterk samen met de score op de Barthel-index ( $r = -0,46$ ), maar ook met die op de GDS ( $r = 0,33$ ) en die op de MMSE ( $r = -0,22$ ).

Aanvullende analyses van de concurrente validiteit bij interviewers die tevens behandelaren waren, lieten zien dat de correlaties over de gehele linie 0,05-0,10 punten sterker waren (cijfers niet gepresenteerd).

**Constructvaliditeit** De constructvaliditeit valt af te lezen uit tabel 4. De aantallen deelnemers per diagnosegroep waren klein (variërend van 4 tot 15 personen), behalve voor affectieve stoornissen. Daarom volstaan we hier met een globale vergelijking van de gemiddelden op de subschalen. Over het algemeen zagen we dat de problemen die met de HONOS65+ gesignaleerd werden, overeenkwamen met de gestelde

diagnose. Zo werden bij mensen met affectieve stoornissen ook problemen geconstateerd bij de subschalen voor andere geestelijke gedragsproblemen, depressieve symptomen en somatische aandoeningen. Ook waren er wat cognitieve problemen en problemen met sociale relaties, wat paste bij de aard van de aandoening. Er waren vrijwel geen problemen op het gebied van alcoholgebruik, opzettelijke zelfverwonding, wanen of hallucinaties bij mensen met affectieve stoornissen. Bij mensen met organische stoornissen lieten de subschalen ADL en somatiek problemen zien, en daarmee samenhangend problemen met dagactiviteiten.

**Gevoeligheid voor verandering** De gevoeligheid voor verandering werd vastgesteld aan de hand van de correlatie tussen verandering op de HONOS65+ en op de BPRS en veranderingen volgens het oordeel van de behandelaar. Veranderingen op de HONOS65+-totaalscore en de -subschalen zijn weergegeven in tabel 5. Bij de totaalscore op de HONOS65+ leken mensen met somatoforme klachten de sterkste verbeteringen te laten zien, maar geen van de waargenomen veranderingen verschilde significant tussen de subgroepen vanwege te kleine aantallen. Verandering in de HONOS65+-totaalscore hing sterk samen met verandering in de BPRS-score ( $r = 0,55$ ;  $p < 0,01$ ), en minder sterk, maar significant, met het oordeel van de behandelaar ( $r = -0,36$ ;  $p <$

TABEL 3 Concurrente validiteit ( $n = 167$ ) van de subschalen van de HONOS65+ (significante correlaties cursief gedrukt)

		MMSE	p	GDS	p	AUDIT	p	BPRS	p	Barthel	p
1	Gedragsstoornissen	0,03	0,71	0,31	0,00	-0,08	0,34	0,43	0,00	0,07	0,37
2	Opzettelijke zelfverwonding	-0,10	0,71	0,49	0,00	-0,08	0,33	0,33	0,00	-0,23	0,00
3	Problemen met alcohol of drugsgebruik	0,18	0,02	-0,03	0,69	0,50	0,00	-0,06	0,48	0,04	0,66
4	Cognitieve problemen	-0,32	0,00	0,18	0,03	-0,11	0,15	0,12	0,13	-0,15	0,05
5	Problemen met somatische ziekten/handicap	-0,26	0,00	0,26	0,00	-0,07	0,36	0,25	0,00	-0,42	0,00
6	Problemen die gepaard gaan met hallucinaties of wanen	0,00	0,96	0,05	0,51	-0,07	0,41	0,20	0,01	-0,22	0,00
7	Problemen met depressieve symptomen	-0,06	0,47	0,78	0,00	-0,18	0,02	0,61	0,00	-0,06	0,42
8	Andere geestelijke en gedragsproblemen	-0,02	0,77	0,28	0,00	-0,10	0,22	0,39	0,00	0,03	0,71
10	Problemen met ADL-activiteiten	-0,22	0,00	0,33	0,00	-0,19	0,01	0,22	0,01	-0,46	0,00

SD = standaarddeviatie, HONOS65+ = Health of the Nations Outcome Scales 65+; MMSE = Mini Mental State Examination, AUDIT = Alcohol Use Disorders Identification Test, BPRS = Brief Psychiatric Rating Scale, ADL = algemene dagelijkse levensverrichtingen.

TABEL 4 Constructvaliditeit van de HONOS65+-subscales en gestelde diagnose door behandelaar

		Affectieve stoornissen		Middelen-misbruik		Hallucinaties/wanen		Somatische stoornissen	
		M	SD	M	SD	M	SD	M	SD
1	Gedragsstoornissen	1,04	(0,90)	0,75	(0,96)	1,00	(1,00)	1,33	(0,98)
2	Opzettelijke zelfverwonding	0,33	(0,98)	0,00	(0,00)	0,27	(0,59)	0,58	(0,90)
3	Problemen met alcohol of drugsgebruik	0,11	(0,39)	1,00	(2,00)	0,13	(0,35)	0,00	(0,00)
4	Cognitieve problemen	0,84	(1,52)	0,50	(0,58)	0,60	(0,63)	0,67	(0,78)
5	Problemen met somatische ziekten/handicap	2,05	(1,01)	1,75	(0,96)	1,93	(1,03)	3,92	(3,18)
6	Problemen die gepaard gaan met hallucinaties of wanen	0,18	(0,60)	0,00	(0,00)	0,93	(2,40)	0,00	(0,00)
7	Problemen met depressieve symptomen	1,81	(1,30)	1,25	(0,50)	1,73	(1,28)	1,92	(1,08)
8	Andere geestelijke en gedragsproblemen	2,37	(2,53)	1,00	(1,15)	2,53	(3,48)	2,75	(2,49)
10	Problemen met ADL-activiteiten	1,32	(1,18)	0,50	(1,00)	1,47	(0,99)	1,17	(1,11)

M = gemiddelde; SD = standaarddeviatie, HONOS65+ = Health of the Nations Outcome Scales 65+, ADL = algemene dagelijkse levensverrichtingen.

0,01). Omdat een hogere score bij de behandelaar duidt op meer verslechtering, was de samenhang negatief.

Het oordeel van de behandelaar hing nog minder sterk samen met veranderingen op de BPRS ( $r = -0,28$ ;  $p < 0,05$ ).

Bij twee bivariate regressieanalyses bleek dat verandering in functioneren volgens de BPRS het best voorspeld werd door veranderingen op de HONOS65+ (33% verklaarde variantie) en in veel mindere mate door het oordeel van de arts (6% verklaarde variantie; cijfers niet gepresenteerd).

## DISCUSSIE

In deze studie zijn de interbeoordelaarsbetrouwbaarheid, de validiteit en de gevoeligheid voor verandering van de HONOS65+ vastgesteld bij 168 ouderen van 60 jaar of ouder die ambulante of intramuraal in behandeling zijn bij een instelling voor geestelijke gezondheidszorg. De resultaten van dit onderzoek beperken zich tot ouderen met affectieve stoornissen, zoals depressie, dysthymie en angst; stoornissen die voorkomen bij 30% van de ouderen die in behandeling zijn bij ggz-instellingen in Nederland (Sectorrapport 2009, GGZ). Ondanks de grootte van de steekproef ble-

TABEL 5 Gemiddelde veranderingen\* op de HONOS65+-subscales naar gestelde diagnose door behandelaar (complete gevallen)

		Affectieve stoornissen (n=69)		Hallucinaties/wanen (n=9)		Somatische stoornissen (n=8)	
		M	SD	M	SD	M	SD
	Totaalscore op de HONOS65+ (alle subscales)	1,64	(4,20)	-1,17	(2,40)	4,17	(5,19)
1	Gedragsstoornissen	0,19	(0,85)	0,44	(1,13)	0,88	(0,83)
2	Opzettelijke zelfverwonding	-0,02	(0,52)	0,11	(0,60)	0,22	(0,67)
3	Problemen met alcohol of drugsgebruik	0,05	(0,31)	-0,22	(0,44)	0,00	(0,00)
4	Cognitieve problemen	0,07	(0,64)	0,33	(0,50)	-0,38	(0,92)
5	Problemen met somatische ziekten/handicap	0,06	(0,91)	-0,38	(0,74)	0,00	(0,63)
6	Problemen die gepaard gaan met hallucinaties of wanen	0,00	(0,64)	-0,50	(1,41)	0,00	(0,00)
7	Problemen met depressieve symptomen	0,53	(1,18)	0,44	(0,53)	0,88	(0,83)
8	Andere geestelijke en gedragsproblemen	0,15	(1,42)	-0,86	(1,21)	0,50	(1,60)
10	Problemen met ADL-activiteiten	0,06	(0,75)	-0,33	(0,71)	0,00	(0,00)

\*Verandering = score  $T_0$  - score  $T_1$ , zodat een positief getal verbetering indiceert. Veranderingen op de HONOS65+ bij ouderen met middenmisbruik zijn weggelaten vanwege het kleine aantal ( $n = 3$ ).

M = gemiddelde; SD = standaarddeviatie; HONOS = Health of the Nations Outcome Scales 65+; ADL = algemene dagelijkse levensverrichtingen.

ken ouderen met verslavingsproblematiek, delier, schizofrenie, persoonlijkheidsstoornissen en slaapstoornissen ondervertegenwoordigd, zodat we geen betrouwbare conclusies konden trekken over de validiteit en de betrouwbaarheid van de HONOS65+ in deze klinische subgroepen.

Van de 9 bestudeerde subschalen geven er 5 een betrouwbare indruk van de ernst van problemen: opzettelijke zelfverwonding, alcoholmisbruik, somatische problemen, depressieve problemen en problemen met ADL-activiteiten. Problemen die een gevolg zijn van gedragsstoornissen zijn redelijk betrouwbaar in kaart te brengen, maar onvoldoende betrouwbaar is een indruk van problemen die samenhangen met cognitieve stoornissen, hallucinaties, en andere psychische ziekten.

De concurrente validiteit was voldoende tot goed; er was een redelijk tot sterke samenhang met schalen die hetzelfde onderliggende construct meten, terwijl deze afwezig was bij schalen die een ander construct meten. Ook de constructvaliditeit was goed. Veelvoorkomende kenmerken bij affectieve stoornissen, zoals depressieve symptomen, somatische problemen, andere geestelijke problemen zoals angst of nervositeit, en problemen met sociale relaties laten op de betreffende subschalen een verhoging van het aantal problemen zien. De aantallen cliënten met andere aandoeningen zijn te klein om daarover uitspraken te doen, maar getallen wijzen wel in de goede richting, zoals een verhoging van de subschaal alcoholmisbruik bij mensen met een alcoholprobleem en een verhoging van de subschaal gedragsstoornissen bij mensen met een somatoforme aandoening. Tot slot stelden we vast dat de HONOS65+ gevoelig is voor verandering.

Onze studie laat in vergelijking met een aantal andere studies op het punt van de betrouwbaarheid een paar opvallende bevindingen zien.

Ten eerste constateren we dat de interbeoordelaarsbetrouwbaarheid in onze studie wat lager uitvalt dan bij sommige andere studies. In de studie van Broersma en Sytema (2008) werden ICC's gevonden van 0,81-0,99. Deze hoge betrouwbaar-

heid is echter gebaseerd op dezelfde informatie van een beperkt aantal casussen, te weten één video-opname van een HONOS65+-interview en twee schriftelijke casussen. In ons onderzoek zijn 31 cliënten afzonderlijk van elkaar geïnterviewd door twee interviewers, waardoor niet alleen de interpretatie van de situatie, maar ook de antwoorden van de cliënt gevarieerd kunnen hebben.

Ook de Franse validatiestudie (Canuto e.a. 2007) laat overwegend betere resultaten zien; de interbeoordelaarsbetrouwbaarheid uitgedrukt in kappa's varieerde van 0,4 (behoorlijk) tot 0,8 (sterk, Landis & Koch 1977). Een belangrijk verschil is echter dat de HONOS65+ in de Franse studie pas afzonderlijk van elkaar werd ingevuld, nadat beoordelaars eerst gezamenlijk de cliënt gesproken hadden en de GDS, de MMSE, en de CAGE hadden afgenomen.

In de studie van Burns e.a. (1999) werd het merendeel (81%) van de interviews voor de betrouwbaarheid afgenomen door twee beoordelaars, wat leidde tot aanzienlijk hogere interbeoordelaarsbetrouwbaarheid van vrijwel alle subschalen.

Een en ander wijst erop dat een goede bekendheid met de cliënt en het gebruik van zoveel mogelijk voorinformatie over het gedrag van de cliënt de betrouwbaarheid van de HONOS65+ aanzienlijk vergroten.

De validiteit, zoals gemeten met de concurrente validiteit en de constructvaliditeit was voldoende tot goed. Twee studies lenen zich voor een vergelijking met onze studieresultaten (Burns e.a. 1999; Canuto e.a. 2007). Onze resultaten wat betreft de validiteit zijn grotendeels in overeenstemming met die van Canuto e.a. (2007). Opmerkelijk was echter dat in onze studie de GDS correleert met subschaal 1 (gedragsstoornissen) en subschaal 2 (opzettelijke zelfverwonding) van de HONOS65+, terwijl in beide andere studies geen samenhang wordt gevonden. Belangrijk verschil is echter dat in onze studie de GDS-30 werd gebruikt, maar in de andere studies de GDS-15. Juist de GDS-30 bevat vragen die in verband gebracht kunnen worden met gedragsproblemen,



zoals vragen over rusteloosheid (item 11) en boosheid (item 24). In onze studie werd verder geconstateerd dat de validiteit van de HONOS65+ groter was bij de interviewers die tevens behandelaren waren.

Voor de praktijk betekent dit dat de waarde van de HONOS65+ als evaluatie-instrument voor de problematiek van oudere cliënten vooral groot is als beoordelaars niet alleen goed getraind, maar ook goed bekend zijn met de cliënt.

Tot slot constateerden we dat de HONOS65+ gevoelig genoeg is om veranderingen te kunnen meten en zelfs een beter instrument is om verandering in kaart te brengen dan het oordeel van de behandelaar alleen. Uiteraard blijft de vraag welk oordeel de klinische situatie het beste benadert, maar als we veronderstellen dat de score op de BPRS de gouden standaard is, dan kan men met de HONOS65+ veranderingen in de ernst van psychiatrische problemen beter vaststellen dan wanneer men zich baseert op het oordeel van de behandelend arts.

**Beperkingen** Onze studie heeft een aantal zwakke kanten die leiden tot een beperkte generaliseerbaarheid van de studieresultaten.

Een eerste beperking betreft de aard van de psychiatrische aandoeningen. Het grootste deel van de cliënten in deze studie had stoornissen van affectieve aard. Hierdoor zijn conclusies over de concurrente validiteit en de constructvaliditeit bij cliënten met hallucinaties of wanen, somatoforme aandoeningen en mensen met alcoholproblemen niet goed te trekken. Hoewel de meest voorkomende aandoening bij cliënten van ggz-instellingen van affectieve aard zijn, is aanvullend onderzoek nodig om de validiteit van de HONOS65+ voor het volledige spectrum van aandoeningen binnen de ouderenpsychiatrie vast te stellen.

Een ander probleem was de relatief grote uitval. Bij bijna een derde ( $n = 52$ ) van de cliënten kon geen vervolgmeting worden afgenomen, voornamelijk doordat ze veranderd waren van behandelaar of niet meer in behandeling waren. Onduidelijk is of dit een relatief gezondere groep ouderen is

omdat verandering van behandelaar of niet meer in behandeling zijn zowel kan wijzen op verslechtering (opname van een extramurale cliënt) als op verbetering (overplaatsing van de kliniek naar de extramurale zorg of ontslag door herstel).

Tot slot constateerden we relatief grote betrouwbaarheidsintervallen rond de puntschatting van de ICC, mogelijk veroorzaakt door verschillen in beoordeling en de relatief kleine aantallen. Hierdoor werd de power om significante correlaties te vinden afgezwakt.

**Sterke kanten** Een sterke kant van de studie is de deelname van meerdere ggz-instellingen verspreid over Nederland. Hierdoor zijn lokale en instellingsafhankelijke invloeden geminimaliseerd. Bovendien is gebruikgemaakt van een breed scala aan valide en betrouwbare meet- en screeningsinstrumenten waarmee de subschalen van de HONOS65+ geëvalueerd konden worden. Dit, en de ruime aantallen ouderen met affectieve stoornissen die deelnamen aan de studie, vergroot het vertrouwen in de conclusies van dit onderzoek.

## CONCLUSIE

De HONOS65+ is een betrouwbaar en valide instrument voor het in kaart brengen van problemen op diverse levensgebieden bij oudere (60+) cliënten met affectieve stoornissen. De betrouwbaarheid en de validiteit van de HONOS65+ worden groter naarmate men over meer informatie van de cliënt beschikt. Dit pleit ervoor de HONOS65+-evaluatie te laten verrichten door behandelaren die goed getraind zijn en die ook andere informatiebronnen dienen te gebruiken, zoals dagelijkse observatie, informatie van de partner of andere verpleegkundigen. Zie hiervoor ook de instructies van de HONOS65+ in Staring e.a. (2003). De HONOS65+ brengt veranderingen goed in kaart en vormt een waardevolle aanvulling op het oordeel van de behandelend artsen. Echter, vervolgonderzoek naar de betrouwbaarheid en de validiteit van de HONOS65+ bij ouderen met

## andere dan affectieve aandoeningen is nodig voor dat de HONOS65+ in de volle breedte van de oude- renpsychiatrie kan worden ingezet.

☞ Dit onderzoek is uitgevoerd onder verantwoordelijkheid van het Nederlands Kenniscentrum Ouderenpsychiatrie in samenwerking met Altrecht GGZ, GGZ-Eindhoven, GGZ-Noord-Holland-Noord en GGZ Westelijk Noord-Brabant. Dit onderzoek was niet mogelijk geweest zonder de medewerking van de deelnemende ouderen.

### LITERATUUR

- Broersma, T.W., & Sytma, S. (2008). Implementatie van het meetinstrument HONOS65+: Onderzoek op de afdeling voor ouderenpsychiatrie. *Tijdschrift voor Psychiatrie*, 50, 77-82.
- Burke, W.J., Houston, M.J., & Boust, S.J. (1989). Use of the Geriatric Depression Scale in dementia of the Alzheimer type. *Journal of the American Geriatrics Society*, 37, 856-860.
- Burns, A., Beevor, A.S., Lelliott, P., e.a. (1999). Health of the Nations Outcome Scales for elderly people (HONOS65+). *The British Journal of Psychiatry*, 174, 424-427.
- Canuto, A., Weber, K., Gold, G., e.a. (2007). Structured assessment of the mental health status in psychogeriatrics: Validity of the French HONOS65+. *The Canadian Journal of Psychiatry*, 52, 37-45.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). 'Mini Mental State': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- GGZ-Nederland (2009). *Zorg op waarde geschat: Sectorrapport GGZ 2009*. Amersfoort, drukkerij Haasbeek.
- Guy, W. (1976). Clinical Global Impressions. In *ECDEU Assessment Manual for Psychopharmacology*. (pp. 218-222). US Department of Health and Human Services.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33, 159-174.
- Mahoney, F.I., & Barthel, D.W. (1965). Functional evaluation: the Barthel Index. *Maryland state medical journal*, 14, 61-65.
- Mulder, C.L., Staring, A.B.P., Loos, J., e.a. (2004). De Health of the Nations Outcome Scales (HONOS) als instrument voor 'routine outcome assessment'. *Tijdschrift voor Psychiatrie*, 46, 273-284.
- Overall, J.E. & Gorham, D.R. (1962). Brief Psychiatric Rating Scale. *Psychological Reports*, 10, 799-812.
- Saunders, J.B., Aasland, O.G., & Babor, T.F. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption-II. *Addiction*, 88, 791-804.
- Shrout, P.E. (1998). Measurement reliability and agreement in psychiatry. *Statistical Methods in Medical Research*, 7, 301-317.
- Shergill, S.S., Shankar, K.L., Seneviratna, K., e.a. (1999). The validity and reliability of the Health of the Nations Outcome Scales (HONOS) in the elderly. *Journal of Mental Health*, 8, 511-521.
- Staring, T., Stobbe, J., & Mulder, N. (2003). HONOS65+ Verklarende tabellen te gebruiken bij de HONOS65+. <http://www.rcpsych.ac.uk/cru/honoscales/honos65/honos65dutch.pdf>.
- Wing, J.K., Beevor, A.S., Curtis, R.H., e.a. (1998). Health of Nations Outcome Scales (HONOS). *British Journal of Psychiatry*, 172, 11-18.
- Yesavage, J.A., Brink, T.L., Lum, O., e.a. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17, 37-49.

### AUTEURS

M.J. AARTSEN is socioloog en gerontoloog bij de Vrije Universiteit, Vakgroep Sociale Wetenschappen, afdeling Sociologie, Amsterdam.

A.K. SPITSBAARD is GZ-psycholoog i.o., Altrecht GGZ, afdeling Indigo, Zeist.

C. VAN BAARSEN is psycholoog, Nederlands Kenniscentrum Ouderen Psychiatrie, Zeist.

A.D.F. DHONDT is psychiater, GGZ-Noord-Holland-Noord, Heiloo.

M. MASCINI is GZ-psycholoog, GGZ-Noord-Holland-Noord, Heiloo.

A. NEFS is gezondheidswetenschapper, GGZ Westelijk Noord-Brabant, Halsteren.

A. SNOEIJERS is nurse practitioner, GGZ Westelijk Noord-Brabant, Halsteren.

M. VAN DIEST is verpleeghuisarts, psychiater, GGZ Westelijk Noord-Brabant, Halsteren.

R. WILTING is psycholoog, GGZ Eindhoven.

Correspondentieadres: dr. M.J. Aartsen, VU, afd. Sociologie, De Boelelaan 1081, 1081 HV Amsterdam.

E-mail: [mj.aartsen@fsw.vu.nl](mailto:mj.aartsen@fsw.vu.nl).

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 3-12-2009.

## SUMMARY

A multi-centre study of the reliability, validity and sensitivity to change of the HONOS65+ in psychiatry for older persons – M.J. Aartsen, A.K. Spitsbaard, C. van Baarsen, A.D.F. Dhondt, M. Mascini, A. Nefs, A. Snoeijers, M. van Diest, R. Wilting –

**BACKGROUND** Within the mental health care services for older persons there is a growing need for insight into and evaluation of the results of clinical treatment. The Health of the Nations Outcome Scales 65+ (HONOS65+) is a promising instrument for the assessment of mental, social and physical health in older persons, but it is not yet known whether it is valid for older persons in the Netherlands.

**AIM** To assess the reliability, validity and sensibility to change of the HONOS65+ when applied to older persons with psychiatric disorders.

**METHOD** The bio-psycho-social level of functioning of clients aged 60 and over (n = 168) receiving mental health care was assessed by means of existing and validated measuring instruments and the results were compared with those obtained with the HONOS65+. Three months later the population sample was re-assessed in order to test the extent to which the HONOS65+ was sensitive to change.

**RESULTS** The reliability and validity of the HONOS65+ could be ascertained for 168 clients aged 60 and over. After three months 116 clients were re-assessed so that the sensitivity of the HONOS65+ to change could be noted.

**CONCLUSION** The HONOS65+ is a reliable and valid instrument for assessing clients with affective disorders such as depression and anxiety and for detecting changes in clients' problems and functioning. No conclusions could be reached regarding the reliability and validity of the HONOS65+ when used for clients with other psychiatric disorders because the clinical subgroups were too small for patterns to be detected.

[TIJDSCHRIFT VOOR PSYCHIATRIE 52(2010)8, 543-553]

**KEY WORDS** HONOS65+, validity study, psychiatry for older persons