

## Betrouwbaarheid van de Zorgzwaartepakket (ZZP) scorelijst

D.H.M. Frijters · W.P. Achterberg

**Abstract The ZZP Questionnaire. Reliability of a new resource utilization measure.** Data to determine the resource utilization of care recipients need to be reliable and the items that are measured need to be useful. In 2006, the Dutch Ministry of Health and Welfare has mandated all nursing homes and homes for the elderly to measure the Resource Utilization of all residents with the ZZP Questionnaire. Are the data resulting from this measurement reliable and is each of the 54 items of the ZZP Questionnaire useful? To answer this we tested the reliability of the data in a nursing home and a home for the elderly in two wards each. For 122 residents questionnaires were completed such that the inter- and intra-rater reliability of the answers could be assessed. Ten of the 54 items in the questionnaire showed insufficient inter rater reliability ( $<0.40$ ) on the weighted Cohen kappa and another sixteen moderate ( $0.40 - 0.60$ ). On the intra rater reliability test seven items had an insufficient kappa and another fifteen moderate. Besides, ten clusters of items could be formed with in-cluster Spearman correlation rates of .75 or higher. From the results of the reliability tests and the item intercorrelation rates we concluded that a substantial number of items needs to be improved and that in the ZZP Questionnaire 15 of the 54 items appear to be redundant on statistical grounds. *Tijdschr Gerontol Geriatr 2007; 38: 166-173*

**Samenvatting** Voor het bepalen van zorgzwaarte van cliënten zijn gegevens nodig die betrouwbaar zijn en die voortkomen uit een doelmatige vragenlijst. Het Ministerie van VWS heeft eind 2006 in alle verpleeg- en

verzorgingshuizen voor elke verblijvende cliënt een scorelijst ZorgZwaartePakketten laten invullen. Is het invullen daarvan betrouwbaar en is het invullen van elk van de 54 vragen noodzakelijk? Om dit te kunnen beantwoorden hebben we onderzoek gedaan naar de betrouwbaarheid van invullen in een verpleeghuis en een verzorgingshuis bij elk twee afdelingen. Voor 122 cliënten werden scorelijsten ingevuld zodanig dat de inter- en de intrabeoordelaarbetrouwbaarheid kon worden beoordeeld. Tien van de 54 vragen in de scorelijst scoorden bij de toets naar interbeoordelaarbetrouwbaarheid onvoldoende ( $<0,40$ ) op de gewogen Cohen kappa overeenkomsttoets en nog eens zestien matig ( $0,40 - 0,60$ ). Ook op intrabeoordelaarbetrouwbaarheid scoorden zeven vragen onvoldoende en nog eens vijftien matig. Daarnaast bleken tien clusters van vragen te kunnen worden gevormd met per cluster een onderlinge Spearman correlatiecoëfficiënt van 0,75 of hoger, hetgeen redundantie van een flink aantal items suggereert. Uit de uitkomsten van de betrouwbaarheidstoetsen en de correlaties tussen de vragen trekken wij de conclusie dat bijna de helft van de vragen in de ZZP scorelijst moet worden verbeterd en dat van de lijst zoals die nu is 15 van de 54 vragen op statistische gronden overbodig lijken.

**Keywords** zorgzwaarte · invulbetrouwbaarheid · verpleeghuis · verzorgingshuis · ZZP scorelijst

### Inleiding

Het Ministerie van Volksgezondheid, Welzijn en Sport (VWS) heeft besloten dat er in 2007 nieuwe regels komen om de kosten voor het verblijf in een zorginstelling te vergoeden.<sup>1</sup> Dit heet zorgzwaartebekostiging en is een

D.H.M. Frijters (✉)  
Afd. Verpleeghuisgeneeskunde, Vrije Universiteit Medisch Centrum, Amsterdam. Tevens: Prismant, Utrecht.

van de einddoelen van de modernisering van de Algemene Wet Bijzondere Ziektekosten (AWBZ). Voor elke zorgvrager wordt berekend op welk zorgzwaartepakket (ZZP) hij of zij recht heeft. Een ZZP is een volledig pakket van zorg dat aansluit op de kenmerken van de cliënt en het soort zorg dat deze nodig heeft. Een ZZP bestaat uit een beschrijving van het type cliënt (een cliëntprofiel), het aantal benodigde uren zorg en een beschrijving van die zorg. Aan ieder ZZP hangt een (maximale) prijs. VWS gaat uit van 45 ZZP's voor de drie sectoren binnen de intramurale zorg; 10 voor de verpleging en verzorging, 7 voor de geestelijke gezondheidszorg en 28 voor de gehandicaptenzorg.

Vooruitlopend op de invoering van de zorgzwaartebekostiging hebben alle instellingen waar intramurale zorg wordt gegeven voor elk van hun cliënten een ZZP scorelijst moeten invullen. Deze scorelijst is door het bureau Hoeksma, Homans & Menting (HHM) in opdracht van en in afstemming met VWS opgesteld. De lijst bevat 54 vragen. Op basis van de antwoorden op deze vragen wordt een cliënt een zorgzwaartepakket (ZZP) toegewezen.

Het invullen van de scorelijsten heeft de instellingen veel tijd gekost en gaat hen in de toekomst bij aanvraag van herindicatie van het zorgzwaartepakket wederom tijd kosten. Voor nieuw op te nemen cliënten zal de scorelijst door het Centrum Indicatiestelling Zorg worden ingevuld. De Nederlandse Zorgautoriteit i.o. (voorheen CTZ) stelt in de toezichttoets Zorgzwaartebekostiging intramurale AWBZ zorg: "Voor een efficiënte vaststelling of AWBZ-uitgaven rechtmatig zijn is het noodzakelijk dat alle partijen in de zorgverlenings- en controleketen kunnen steunen op voldoende en betrouwbare gegevens..."<sup>2</sup> Noch in de informatie op de VWS-website, noch in de (inter)nationale literatuur konden wij echter gegevens vinden over de betrouwbaarheid, validiteit, responsiviteit of de doelmatigheid van de vragenlijst. In verschillende andere landen in Europa en Noord-Amerika zijn wel zorgzwaartebekostigingssystemen voor de chronische zorg uitgebreid onderzocht en in de literatuur beschreven.<sup>3-9</sup>

Het hier gepresenteerde onderzoek wil een antwoord geven op de volgende twee vragen:

- (1) Wat is de betrouwbaarheid van de vragen van de ZZP scorelijst?
- (2) Welke vragen hebben statistisch sterke overeenkomst met elkaar?

De laatste vraag is belangrijk voor de doelmatigheid van de scorelijst, omdat een sterke correlatie tussen scores op verschillende vragen op itemredundantie kan wijzen.

## Methoden

### *Dataverzameling*

Twee instellingen, verpleeghuis Het Zonnehuis Zuidhorn en verzorgingshuis Drieënhuizen, namen deel aan het onderzoek met ieder twee afdelingen. Op de deelnemende afdelingen werden voor alle verblijvende cliënten als volgt de scorelijsten ingevuld.

Op de vier afdelingen (van gemiddeld ruim 30 cliënten) zijn in november/december 2006 telkens door één beoordelaar per afdeling voor alle cliënten van die afdeling in een periode van een aantal weken alle ZZP scorelijsten ingevuld. Van deze invullingen werd een papieren kopie gemaakt. In de tweede en derde week van februari 2007 werd in het kader van het hier beschreven onderzoek door dezelfde beoordelaar uit de november/december periode voor de nog verblijvende cliënten op papier wederom de ZZP scorelijst ingevuld (voor bepalen van de intrabeoordelaarbetrouwbaarheid). Door andere beoordelaars (twee per afdeling) werd onafhankelijk daarvan voor elk van de nog verblijvende cliënten ook op papier een ZZP scorelijst ingevuld (voor bepalen van de interbeoordelaarbetrouwbaarheid).

Voor de beoordeling van de interbeoordelaarbetrouwbaarheid werd de vergelijking van de tweede beoordeling van de eerste beoordelaar en die van de nieuwe beoordelaar gebruikt. Elk van de beoordelaars was werkzaam op de afdeling, had een opleiding Verzorgende Individuele Gezondheidszorg (VIG) op niveau 3 of hoger en kende de cliënt goed. De beoordelaars hadden buiten de instelling in het kader van het VWS ZZP project dan wel in de instelling een uitvoerige instructie en training ontvangen.

De ingevulde scorelijsten zijn geanonimiseerd aan de onderzoekers ter beschikking gesteld.

### *ZZP scorelijst*

De Zorgzwaartepakket scorelijst bevat vragen over 54 onderwerpen, gerangschikt naar begeleidingsbehoefte (vragen 1 t/m 22); verzorgingsbehoefte (vragen 23 t/m 34); verplegingsbehoefte (vragen 35 t/m 37); behoefte aan gedragsregulering (vragen 38 t/m 46); cognitieve/psychische aspecten (vragen 47 t/m 52); en benodigde setting van zorgverlening (vragen 53 en 54). Alle vragen hebben een oplopende score, wat wil zeggen dat '0' een lichtere conditie weergeeft dan '1', dan '2' enzovoort. Op vijf vragen na zijn de scoremogelijkheden telkens 0–3, waarbij 0 en 3 de uitersten 'geen hulp' en 'volledige/altijd hulp' en 1 en 2 de tussenwaarden 'enige/soms' en 'vaak/veel hulp' weergeven. Twee vragen hebben een scorebereik van 1–4, één van 1–3, één van 0–2 en één van 0–1.

*Statistische analyse*

De betrouwbaarheid van de ZZP scorelijst werd beoordeeld met de kappaa-overeenstemmingstoets.<sup>10,11</sup> De kappaa heeft een waardebereik van -1 tot +1. Een waarde 0 betekent “geen overeenkomst, m.a.w. toeval”, een waarde +1 “volledig gelijk” en -1 “volledig ongelijk”. Indien het aantal scoremogelijkheden van een variabele groter was dan 2, werd een kwadratisch gewogen kappaa gemeten.<sup>12</sup> De kappawaarde is voor dit soort toetsen bijzonder bruikbaar omdat de toevallige kans op overeenstemming wordt meegewogen.

Voor de interpretatie van de kappawaarde maakten wij gebruik van de Koch en Landis kwalificatie: een kappawaarde groter dan 0,80 is uitstekend, van 0,60 tot 0,80 redelijk tot goed, van 0,40 tot 0,60 matig en lager dan 0,40 onvoldoende.<sup>13</sup>

*Overeenkomsten tussen de vragen*

De overeenkomst in de antwoorden tussen de verschillende vragen van de scorelijst werd onderzocht met de Spearman correlatietoets (SPSS versie 11). Wanneer binnen hetzelfde thema (begeleidingsbehoefte, verzorgingsbehoefte, verplegingsbehoefte, behoefte aan gedragsregulering, cognitieve/psychische aspecten, benodigde setting van zorgverlening) antwoorden op twee verschillende vragen voor tenminste 0,75 met elkaar correleerden, werden deze in een cluster geplaatst. Was er nog een vraag binnen het thema dat met elk van de twee al in het cluster geplaatste vragen voor tenminste 0,75 correleerde, dan werd ook deze vraag aan het cluster toegevoegd enzovoort.

**Resultaten***Dataverzameling*

In november/december 2006 hadden vier beoordelaars op de vier afdelingen die aan het onderzoek zouden gaan deelnemen, van alle verblijvende cliënten de ZZP scorelijsten ingevuld met behulp van de ZZP scorelijst Internetmodule van VWS.

Voor alle cliënten (122) die eind januari 2007 nog op de vier afdelingen van de twee instellingen verbleven werd op dat moment opnieuw een scorelijst ingevuld. De volledigheid van invulling van de vragen van de scorelijsten was in november/december 100% en in januari bij de meeste vragen wederom 100%, maar bij elk vraag tenminste hoger dan 98%. In totaal werden door twaalf verpleegkundigen ZZP scorelijsten ingevuld.

*Betrouwbaarheid**Interbeoordelaarbetrouwbaarheid*

Tabel 1 laat zien dat tien van de 54 vragen van de scorelijst een onvoldoende overeenstemming (kappaa < 0,40) tussen beoordelaars hadden. Het gaat om de vragen 35, 36, 37, 38, 43, 45, 49, 50, 52 en 54 die handelen over wond/stoma zorg, zorg bij uitscheiding, gespecialiseerd verpleegkundig handelen, bescherming tegen zelfverwonding, dwangmatig gedrag, ontremd gedrag, vertekend omgevingsbeeld, motivatie, persoonlijke relaties en type verblijfsomgeving. Zestien vragen stemden matig overeen (kappaa > 0,40 en < 0,60), vierentwintig vragen stemden redelijk tot goed overeen (kappaa > 0,60 en < 0,80) en vier stemden uitstekend overeen (kappaa > 0,80).

*Intrabeoordelaarbetrouwbaarheid*

Een tweede beoordeling ten opzichte van de eerste door dezelfde beoordelaar stemde over het algemeen wat beter overeen dan scores van twee verschillende beoordelaars. Zeven vragen stemden onvoldoende overeen (kappaa < 0,40), waarvan vijf die dat ook al deden tussen beoordelaars. Negentien vragen stemden matig overeen (kappaa > 0,40 en < 0,60), zesentwintig redelijk tot goed (kappaa > 0,60 en < 0,80) en twee stemden uitstekend overeen (kappaa > 0,80).

De overeenstemmingsresultaten waren bij het verpleeghuis over de gehele linie statistisch significant hoger dan in het verzorgingshuis.

*Overeenstemming in scores op de vragen van de ZZP scorelijst*

Tabel 2 laat zien dat met de resultaten van de Spearman correlatietoets in totaal tien clusters van (inhoudelijk samenhangende) vragen kunnen worden gevormd die voor tenminste 0,75 in hun antwoorden met elkaar correleerden en zeven clusters die zelfs voor tenminste 0,80 met elkaar correleerden. Dit was in alle gevallen statistisch significant met een P-waarde van < 0,001 (tweezijdig getoetst).

**Discussie**

Niet eerder is er een wetenschappelijk verslag verschenen over de eigenschappen van de methode om te komen tot zorgzwaartebekostiging in de Verpleeg- en Verzorgingssector (V&V) via de ZZP methode. Dit onderzoek bij 122 verpleeg- en verzorgingshuiscliënten laat zien dat zowel de inter- als de intrabeoordelaarbetrouwbaarheid van dit systeem te wensen overlaat: tien van de 54 vragen hebben

**Tabel 1** ZZZ scorelijst: inter- en intrabeoordelaarbetrouwbaarheid (kappa met 95% betrouwbaarheidsinterval (B.I.), vergelijking tussen twee locaties.

Vraag	Inter		3H/HZZ **	Intra		3H/HZZ **
	kappa	95% B.I.		kappa	95% B.I.	
V1 Oriëntatie in tijd	0,66	(,54 – ,77)	-	0,74	(,66 – ,82)	+
V2 Oriëntatie in plaats	0,74	(,64 – ,85)	-	0,78	(,69 – ,87)	
V3 Oriëntatie in ruimte	0,66	(,55 – ,77)	-	0,80	(,72 – ,89)	+
V4 Oriëntatie in persoon	0,60	(,48 – ,73)	-	0,66	(,55 – ,77)	-
V5 Beslisbekwaamheid	0,65	(,54 – ,75)	+	0,55	(,39 – ,70)	+
V6 Uitvoeren eenvoudige taken	0,56	(,42 – ,70)	-	0,53	(,38 – ,69)	-
V7 Uitvoeren complexere taken	0,59	(,41 – ,78)	-	0,44	(,20 – ,67)	-
V8 Dagelijkse routine regelen	0,63	(,51 – ,75)	-	0,65	(,53 – ,76)	+
V9 Uitvoeren dagelijkse bezigheden	0,72	(,61 – ,83)	-	0,61	(,20 – ,67)	-
V10 Geld beheren, kopen, betalen	0,75	(,64 – ,85)	+	0,75	(,48 – ,74)	+
V11 Afhandelen administratie	0,76	(,65 – ,87)	-	0,50	(,65 – ,86)	-
V12 Openbaar vervoer gebruiken	0,67	(,47 – ,86)	-	0,32	(-,09 – ,73)	+
V13 Zich redden in winkels, bank	0,78	(,70 – ,86)	-	0,65	(,49 – ,81)	+
V14 Anderen begrijpen	0,66	(,56 – ,76)	-	0,64	(,55 – ,73)	-
V15 Zich begrijpelijk maken	0,72	(,63 – ,80)	-	0,70	(,60 – ,79)	-
V16 Gesprek voeren	0,62	(,50 – ,74)	-	0,66	(,56 – ,76)	-
V17 Contacten leggen	0,62	(,50 – ,75)	-	0,62	(,51 – ,74)	-
V18 Deelnemen aan clubs	0,66	(,54 – ,79)	+	0,43	(,23 – ,62)	-
V19 Vrije tijd buitenshuis	0,67	(,54 – ,81)	-	0,42	(,20 – ,64)	-
V20 Bezoeken algemene voorzieningen	0,60	(,46 – ,74)	-	0,63	(,47 – ,79)	+
V21 Dagtaak uitvoeren	0,70	(,59 – ,81)	+	0,65	(,53 – ,77)	-
V22 Centrale doel van begeleiding	0,53	(,39 – ,68)	-	0,38	(,21 – ,55)	-
V23 Zich wassen	0,67	(,56 – ,78)	-	0,70	(,58 – ,81)	-
V24 Zich aan- en uitkleden	0,72	(,62 – ,82)	-	0,78	(,69 – ,87)	-
V25 In en uit bed stappen	0,82	(,75 – ,90)	-	0,83	(,76 – ,89)	-
V26 Toilet bezoeken	0,84	(,77 – ,91)	-	0,87	(,82 – ,92)	-
V27 Eten en drinken	0,51	(,39 – ,64)	-	0,62	(,50 – ,74)	-
V28 Zelfverzorging	0,55	(,41 – ,69)	-	0,55	(,41 – ,68)	+
V29 In zit- of lighouding komen	0,83	(,76 – ,91)	-	0,76	(,65 – ,86)	-
V30 Lichaamspositie veranderen	0,83	(,77 – ,90)	-	0,75	(,66 – ,83)	-
V31 Binnenshuis voortbewegen	0,74	(,63 – ,85)	+	0,75	(,64 – ,86)	-
V32 Korte afstanden lopen	0,62	(,50 – ,75)	-	0,53	(,38 – ,67)	-
V33 Gebruik rechter- of linkerhand	0,66	(,52 – ,79)	-	0,72	(,59 – ,84)	-
V34 Grove hand- en armbewegingen	0,61	(,50 – ,73)	-	0,60	(,50 – ,71)	-
V35 Wond- en /of stomaverzorging	0,33	(,11 – ,56)	+	0,28	(,05 – ,51)	-
V36 Ondersteuning bij uitscheiding	0,33	(,06 – ,61)	-	0,32	(,04 – ,61)	-
V37 Gespecialiseerd verplk. handelen	0,33	(-,15 – ,81)	-	0,33	(-,15 – ,85)	-
V38 Zelfverwonding	0,20	(-,24 – ,65)	-	0,43	(,06 – ,81)	-
V39 Verbaal agressief gedrag	0,53	(,30 – ,79)	-	0,72	(,60 – ,84)	-
V40 Lichamelijk agressief gedrag	0,46	(,06 – ,86)	-	0,73	(,57 – ,88)	-
V41 Destructief gedrag	0,49	(-,20 – 1,00)	-	0,22	(-,76 – 1,00)	?
V42 Manipulatief gedrag	0,54	(,35 – ,74)	-	0,56	(,35 – ,76)	-
V43 Dwangmatig gedrag	0,33	(,10 – ,55)	-	0,54	(,38 – ,71)	-
V44 Ongepast seksueel gedrag	0,62	(,17 – 1,00)	-	0,65	(,36 – ,95)	-
V45 Ontremd gedrag	0,25	(-,05 – ,55)	-	0,52	(,30 – ,73)	-
V46 Psychiatrische problematiek	0,49	(,27 – ,72)	-	0,60	(,44 – ,77)	-
V47 Concentratiestoornissen	0,48	(,33 – ,63)	+	0,54	(,42 – ,66)	-

**Tabel 1** (continued)

Vraag	Inter		3H/HZZ <sup>**</sup>	Intra		3H/HZZ <sup>**</sup>
	kappa	95% B.I.		kappa	95% B.I.	
V48 Geheugen- en denkstoornissen	0,57	(,43 – ,71)	+	0,64	(,53 – ,75)	+
V49 Vertekend omgevingsbeeld	0,30	(,10 – ,50)	-	0,54	(,37 – ,71)	-
V50 Onvoldoende motivatie	0,38	(,23 – ,53)	+	0,59	(,47 – ,70)	+
V51 Angststoornis	0,47	(,31 – ,64)	+	0,41	(,22 – ,60)	+
V52 Gemis persoonlijke relaties	0,36	(,18 – ,54)	+	0,19	(-,05 – ,43)	-
V53 Leveringsvoorwaarde zorg	0,51	(,35 – ,66)	-	0,55	(,46 – ,65)	-
V54 Type verblijfsomgeving	0,40	(,24 – ,57)	-	0,74	(,64 – ,85)	-

<sup>\*\*</sup> De vergelijking van de kappacoëfficiënt in verzorgingshuis 3H ten opzichte van verpleeghuis HZZ, waarbij + een hogere kappa betekent voor het verzorgingshuis, – een lagere en ? niet te bepalen.

een te lage interbeoordelaarbetrouwbaarheid (kappa < 0,40) en nog eens zestien scores hierop matig. De indruk bestaat dat lastiger te interpreteren items (motivatie, zelfverwonding, gedrag) slechter scoren dan items over transfers (in zit of lighouding komen, lichaamspositie veranderen, in en uit bed stappen, toiletbezoek) die uitstekend (kappa >0,80) scoren. Ook voor de intrabeoordelaarbetrouwbaarheid geldt dat een groot deel van de vragen (16/54) onvoldoende tot matig scoorde. Een bijkomende reden voor de matige betrouwbaarheid hiervan zou kunnen liggen in mogelijke veranderingen in de toestand van de cliënten in de relatief lange tijdsspanne (circa zes weken) tussen de eerste en tweede beoordeling. Daarentegen zouden wij dat effect van die veranderingen dan zeker ook verwacht hebben bij de items die zowel bij de inter- als bij de intrabeoordelaarsbetrouwbaarheid uitstekend scoorden.

ZZP scorelijsten worden behalve door ziekenverzorgenden in verpleeg- en verzorgingshuizen ook ingevuld door medewerkers van de thuiszorg, medewerkers bij

instellingen voor verstandelijk gehandicapten en, niet te vergeten, door medewerkers bij het Centrum Indicatiestelling Zorg (CIZ). Het lijkt niet waarschijnlijk dat vragen die in ons onderzoek een slechte betrouwbaarheid hadden in andere settings beter zullen presteren. De beoordelaars in het verpleeghuis en het verzorgingshuis waren goed opgeleid, kenden hun cliënten vanuit de dagelijkse omgang en waren goed geschoold voor het invullen van de ZZP scorelijst. Beoordelaars elders, met name bij het CIZ, zullen vaak onder minder gunstige omstandigheden hun gegevens moeten verzamelen. Omdat de CIZ-medewerkers een centrale positie gaan krijgen bij de toewijzing van de ZZP's is aanvullend onderzoek naar betrouwbaarheid in deze setting dan ook zeker wenselijk.

Daarnaast vonden wij statistische aanwijzingen voor mogelijke redundantie van een flink aantal vragen. Zo konden wij tien clusters van vragen samenstellen die sterk met elkaar zijn gecorreleerd. De mogelijkheid zou onderzocht moeten worden of deze clusters tot telkens één

**Tabel 2** Spearman correlatie in antwoorden tussen vragen van de ZZP scorelijst (N = 122).

ZZP-scorelijst vragen	inhoud vragen	laagste Spearman correlatie:> 0,75 (+); > 0,80 (++)*
1	1, 2, 3 oriëntatie in tijd, plaats en ruimte	++
2	6, 7 eenvoudige en complexe taken kunnen uitvoeren	+
3	8, 9 dagelijkse routine regelen en uitvoeren	++
4	11, 13 administratieve zaken afhandelen en zich in winkel en bank kunnen redden	++
5	14, 15 anderen kunnen begrijpen en zichzelf begrijpelijk maken	++
6a	18,19, 20 deelname aan clubs, vrije tijd buitenshuis en bezoek aan algemene voorzieningen	+
6b	19, 20	++
7a	23, 24, 28 wassen, kleden, zelfverzorging	+
7b	23, 24	++
8	25, 29, 30, 31 in/uit bed komen, kunnen gaan zitten, lichaamspositie, zich voortbewegen	++
9	39, 40 verbaal en lichamelijk agressief gedrag	+
10	47, 48 concentratie en geheugen	+

\* Alle correlaties hebben een P<0,001 (tweezijdig getoetst).

vraag te herformuleren zijn, waardoor het aantal vragen sterk kan worden beperkt. Dit geeft mogelijk een besparing van 15 vragen. Het betreft clusters van vragen over oriëntatie, kunnen uitvoeren van taken, de dagelijkse routine, regelen van administratieve zaken, activiteiten buitenshuis, ADL zelfzorg, lichamelijk functioneren, agressief gedrag en denkvermogen.

We hebben in dit onderzoek weliswaar de betrouwbaarheid en doelmatigheid van de (hoeveelheid) vragen gemeten, maar kunnen daarmee geen uitspraak doen over de validiteit van de ZZP of de betrouwbaarheid van toewijzing van Zorgzwaartepakketten aan cliënten. Dit kunnen we niet omdat het algoritme waarmee de ZZP toewijzing wordt berekend niet is vrijgegeven. Mogelijk dat door een samennemen van scoremogelijkheden op een vraag, bijvoorbeeld 0 en 1 samen en 2 en 3 samen de betreffende vraag een hogere kappa bij de overeenstemmingstoets laat zien en daarmee als vraag toch geschikt is voor de ZZP toewijzing. Voor de vragen die sterk met elkaar correleren blijft echter het feit dat ze, hoe het algoritme ook is opgebouwd, zonder wezenlijk verlies aan informatie kunnen worden ingedikt of kunnen vervallen. Omdat het hier per jaar om honderdduizenden vragenlijsten gaat is wezenlijke efficiëntiewinst te behalen. Een nog beter inzicht in de kwaliteiten van de ZZP zou verkregen kunnen worden door aanvullend onderzoek waarbij schaalanalysemethoden volgens de moderne item-responstheorie worden gebruikt.

Een vraag die naar ons weten tot nu ten aanzien van de ZZP scorelijst nergens volhardend is gesteld, maar die wellicht de belangrijkste van alle is, is of de ZZP scorelijst bij gebruik als onderliggend systeem voor kostenvergoeding de druk van majoreren kan doorstaan. Majoreren betekent dat de zorgverlener primair geleid wordt door mogelijke effecten op zijn inkomsten. De Nederlandse Zorgautoriteit (NZA) gebruikt hiervoor de term ‘upcoding’, ofwel het koppelen van een cliënt aan een duurdere ZZP dan nodig gelet op zijn zorgbehoefte. Uit het buitenland is bekend (VS, Canada) dat een systeem dat los van de praktijk van de zorgverlening gegevens verzamelt voor kostenvergoeding gevoelig is voor majoreren en veel controle nodig heeft van buitenaf. Om die reden zijn in veel landen kostenvergoedingssystemen ontwikkeld die gebaseerd zijn op gegevensvastleggingen die primair bedoeld zijn om de zorgverlening te ondersteunen en die ingebouwde controles hebben in de vorm van ‘grote redundantie’ en ontmoedigen van onterecht hoge scores doordat daarmee tevens slecht op indicatoren van kwaliteit van zorg wordt gescoord.<sup>4,5,8,9,14</sup> Naast aanvullend onderzoek naar de betrouwbaarheid van de vragenlijst, zal validerend onderzoek met reeds geteste zorgzwaarte-instrumenten nodig zijn om deze grote verandering in de bekostiging van de V&V sector met wetenschappelijke

argumenten te staven. Zo zijn er nog vele vragen onbeantwoord. Is de vragenlijst bijvoorbeeld in staat om belangrijke veranderingen in de zorgvraag van de patiënt aan te geven? Geeft het systeem mogelijk prikkels om extra activiteiten die de kwaliteit van leven ten goede komen, zoals snoezelen of Passiviteiten Dagelijks Leven (PDL), niet uit te voeren? Wat is de invloed op zorgactiviteiten, zoals het bevorderen van een grotere zelfstandigheid of autonomie, die kunnen leiden tot een lager ZZP en een lagere bekostiging?

Samenvattend geeft dit onderzoek aanleiding om de betrouwbaarheid en doelmatigheid van de ZZP scorelijst in twijfel te trekken. Aanvullend onderzoek, met een grotere populatie en in verschillende settings, is daarom noodzakelijk. Daarnaast vragen we ons af hoe verstandig het is om de zorgbekostiging te baseren op een systeem met een zo fragiele wetenschappelijke basis als de ZZP Scorelijst.

## Dankwoord

We danken Tom van der Meulen, voorzitter van de Raad van Bestuur van Zorgcombinatie Nieuwe Maas, en Jan Peter Kroes, voorzitter van Woonzorgcentra Westerkwartier voor het faciliteren van het onderzoek bij verzorgingshuis Drieën-Huijsen en verpleeghuis Het Zonnehuis Zuidhorn. Tevens danken we de medewerkers van het verzorgingshuis en het verpleeghuis die de gegevens hebben verzameld en die de verzameling hebben voorbereid en georganiseerd.

## Literatuur

- Website VWS: [www.MinVWS.nl](http://www.MinVWS.nl), 2006.
- Toezichttoets Zorgzwaarte bekostiging Intramurale AWBZ zorg. [www.ggzbeleid.nl/pdfdbc/fb\\_zs\\_toetszzps\\_0609.pdf](http://www.ggzbeleid.nl/pdfdbc/fb_zs_toetszzps_0609.pdf)
- Hawes C, Morris JN, Phillips CD, Mor V, Fries BE, Nonemaker S. Reliability estimates for the Minimum Data Set for nursing home residents assessments and care screening (MDS). *Gerontologist* 1995;35:172-8.
- Frijters DHM, Achterberg WP, Hirdes JP, Fries BE, Morris JN, Steel K. Geïntegreerd gezondheidsinformatiesysteem op basis van Resident Assessment Instrumenten. *Tijdschr Gerontol Geriatr* 2001; 32: 8-16.
- Fries BE, Schneider DP, Foley WJ, Gavazzi M, Burke R, Cornelius E. Refining a case-mix measure for nursing homes: Resource Utilization Groups (RUG-III). *Medical Care* 32 (7) 1994: 668-85.
- Carpenter GI, Ikegami N, Ljunggren G, Carrillo E, Fries BE. RUG-III and resource allocation: comparing the relationship of direct care time with patient characteristics in five countries. *Age Ageing*. 1997 Sep;26 Suppl 2:61-5.
- Bjorkgren MA, Hakkinen U, Finne-Soveri UH, Fries BE. Validity and reliability of Resource Utilization Groups (RUG-III) in Finnish long-term care facilities. *Scand J Public Health*. 1999 Sep;27(3):228-34.

- Carpenter GI, Main A, Turner GF. Casemix for the elderly inpatient: Resource Utilization Groups (RUGs) validation project. *Age Ageing* (24) 1995: 5 -13.
- Bjorkgren MA, Fries BE, Shugarman LR. A RUG-III Case-Mix System for Home Care. *Canadian Journal on Aging* (19) 2000: 106-25.
- Fleiss JL. *Statistical methods for rates and proportions*, 2nd Ed. New York: John Wiley & Sons, 1981.
- Cohen J. Weighted kappa: nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol Bull* 1968;70:213-20.
- Frijters DHM, Gerritsen DL, Steverink N. Zorginhoudelijke kwaliteit: betrouwbaarheid en bruikbaarheid van observatiegegevens in de benchmark voor verpleeg- en verzorgingshuizen. *Tijdschr Gerontol Geriatr* 2003; 34: 21-9.
- Landis JR, Koch GC. A review of statistical methods in the analysis of data arising from observer reliability studies (part 1). *Statistica Neerlandica* 1975; 29: 101-123.
- Mor V, Berg K, Angelelli J, Gifford D, Morris J, Moore T. The Quality of Quality Measurement in U.S. Nursing Homes. *Gerontologist* (43 Spec. No 2) 2003: 37-46.