

Het meten van kwaliteit van leven in het verpleeghuis

Dissertaties worden besproken en gepromoveerden kunnen reageren.

Het meten van kwaliteit van leven in het verpleeghuis

Debby Gerritsen: *Quality of life and its measurement in nursing homes*

Promotie: 17 september 2004, Vrije Universiteit Amsterdam

Promotor: Prof.dr. M.W. Ribbe

Bespreking: Dr. S. Teunisse, neuropsycholoog, ouder- en psycholoog, (ere)docent Universiteit van Aberdeen

Kwaliteit van leven is een van de belangrijkste uitkomstmaten van medisch handelen bij chronische aandoeningen. Ook binnen de verpleeghuis sector wordt kwaliteit van leven steeds vaker gezien als een belangrijke uitkomstmaat voor de kwaliteit van zorg. Dit proefschrift sluit aan bij deze ontwikkeling en richt zich op de vragen wat er wordt verstaan onder kwaliteit van leven in het verpleeghuis en hoe het in kaart gebracht kan worden bij bewoners die vaak niet meer in staat zijn te communiceren over hun gevoelens, wensen en behoeften. Het fraaie van dit - prachtig vormgegeven - proefschrift is dat eerst een theorie over kwaliteit van leven in het verpleeghuis wordt geformuleerd en dat meetinstrumentarium zoals gebruikt in de praktijk, daaraan vervolgens worden getoetst: de onderzoeker evalueert of een observatie methode zoals gebruikt in verpleeghuizen (de 'Minimum Data Set' (MDS) van de 'Resident Assessment Instrument' (RAI)) volstaat als instrument voor het meten van kwaliteit van leven.

Na een inleidend hoofdstuk (hoofdstuk 1) over de context van dit onderzoek en de vraagstellingen, wordt de zoektocht naar een passend model voor kwaliteit van leven in het verpleeghuis beschreven (hoofdstuk 2). De literatuur wordt met behulp van zoektermen nagezocht op modellen van kwaliteit van leven, die voldoen aan verschillende criteria: het model dient universeel en breed geformuleerd (d.w.z. niet alleen gericht op

gezondheid) te zijn, informatie te geven over de onderliggende dimensies en hoe deze met elkaar samenhangen en elkaar beïnvloeden, en moet tenslotte ruimte bieden aan individuele verschillen in de beleving van kwaliteit van leven. Van de zes modellen die in de literatuur worden gevonden voldoet alleen de Social Production Functions (SPF) theorie aan alle criteria. Het betreft een hiërarchisch model waarbij algeheel welbevinden (het hoogste niveau) voortkomt uit het gebruik en de (her) waardering van persoonlijke bronnen (het laagste niveau). Bronnen zijn dingen die iemand heeft of doet, zoals het hebben van een partner, vrienden, zelfstandig functioneren, en cognitief functioneren. Verlies van de ene bron kan tot op zekere hoogte gecompenseerd kan worden door het gebruiken of herwaarderen van andere bronnen. Een interessant model dat aansluit bij recente ontwikkelingen in de algemene kwaliteit van leven literatuur, waar individuele verschillen in beleving en compensatie mechanismen steeds meer aandacht krijgen.

Alle volgende hoofdstukken van het proefschrift richten zich op het vinden van passend meetinstrumentarium voor dit model. Daarbij wordt in eerste instantie uitgegaan van de bestaande MDS-RAI observatie schalen (hoofdstukken 3-4), maar waar nodig worden nieuwe schalen geconstrueerd (hoofdstukken 5-6) of andere bestaande schalen geëvalueerd (hoofdstuk 7). Metingen vinden plaats bij verpleeghuizen, die de RAI reeds hebben ingevoerd als geautomatiseerd cliëntvolgsysteem voor het in kaart brengen van medische, functionele, psychosociale en cognitieve toestand van de bewoners. Deze verpleeghuizen wordt gevraagd de MDS-RAI schalen te hanteren zoals gebruikelijk in de dagelijkse praktijk. Een indrukwekkende hoeveelheid data is verzameld. Enkelvoudige metingen zijn beschikbaar voor tenminste 227 (Nederlandse) bewoners, soms oplopend tot 1909

bewoners (data verkregen via een Canadese studie); herhaalde metingen in het kader van test-hertest en interbeoordelaarsbetrouwbaarheid studies zijn beschikbaar voor ongeveer 150 bewoners. Psychometrische analyses richten zich op interne betrouwbaarheid, test-hertest betrouwbaarheid, interbeoordelaars betrouwbaarheid, inhoudsvaliditeit (aan de hand van literatuurstudie en met behulp van een panel van experts), begripsvaliditeit (aan de hand van convergentie met andere maten en factor analyse) en waar mogelijk worden bevindingen zoals verkregen in de ene steekproef gechecked in een andere steekproef (cross-validatie). Tenslotte worden bevindingen verkregen bij somatische en psychogeriatrische bewoners met elkaar vergeleken, evenals bevindingen van bewoners met lichte en ernstige cognitieve achteruitgang. Al met al een uitgebreide, diverse en gedegen psychometrische evaluatie, waarbij slechts één kleine kritische noot, namelijk dat bij validatie soms gebruik gemaakt wordt van vergelijkingsschalen die niet optimaal valide zijn volgens de bevindingen in andere hoofdstukken (bijvoorbeeld de Depressie Lijst als vergelijkingsschaal voor de Depressie Rating Schaal waarbij bewoners met ernstige cognitieve achteruitgang niet worden uitgesloten, en de Depressie Rating Schaal als vergelijkingsschaal voor de Sociale Betrokkenheid schaal). Doordat de gegevens voor de MDS-RAI schalen worden verzameld zoals gangbaar in de praktijk van alledag (dus zonder additionele training van de verpleegkundigen of zonder inschakeling van onderzoekersverpleegkundigen) waar niet alles volgens de regels verloopt, merkt de onderzoeker terecht op dat de vermelde betrouwbaarheden en andere waarden een ondergrens weergeven.

De onderzoeker concludeert dat de MDS-RAI schalen vooral gericht zijn op het in kaart brengen van de bronnen in het SPF model en dat de ADL, de cognitieve schaal en de aangepaste versie van de Sociale Betrokkenheid schaal van de MDS daarvoor goed bruikbaar zijn. Test-hertest en interbeoordelaars betrouwbaarheid van de MDS-depressieschaal daarentegen zijn minder dan die van de andere schalen. Volgens de onderzoeker komt dit doordat psychologische aspecten van gedrag moeilijker te observeren zijn. De betrouwbaarheden verkregen in deze studie zijn gebaseerd op een design waarbij test en hertest meting, evenals de metingen van de verschillende observatoren binnen 1 maand na elkaar plaatsvinden. Het lijkt niet onmogelijk dat stemming meer varieert over tijd dan cognitieve en ADL functioneren. Daarom de vraag of het ook mogelijk is dat de gevonden lagere betrouwbaarheden voor de MDS-depressie schaal juist een indicatie zijn van goede responsiviteit en dus validiteit?

De onderzoeker concludeert tevens dat de MDS-RAI weinig aandacht besteed aan welbevinden, een centraal

concept in het SPF-model, en dat additionele maten nodig zijn voor het meten van sociaal en lichamelijk welbevinden, evenals voor het algeheel subjectief welbevinden. Daartoe ontwikkelt zij met succes twee nieuwe schalen voor sociaal welbevinden (het Prikkelend Gedrag Profiel en de 'Social Well-Being Of Nursing home residents' schaal). Ook vergelijkt zij de waarde van zes bestaande zelfbeoordelingsschalen voor algeheel welbevinden. Dit laatste is een interessant exercitie omdat er vaak van wordt uitgegaan dat verpleeghuisbewoners niet in staat zijn tot het beantwoorden van dergelijke zelfbeoordelingsvragenlijsten. Toch blijkt dat van de bewoners met een MMSE van 5 of hoger, allen in staat zijn tot het beantwoorden van de vragen uit de Depressie Lijst en dat het merendeel ook in staat is de vragen uit de Geriatric Depression Scale en de Philadelphia Geriatric Center Morale Scale te beantwoorden. De Depressie Lijst, van origine ontwikkeld voor mensen met dementie, heeft samen met de Geriatrische Depressie Schaal de beste betrouwbaarheid en validiteit. De bevindingen zijn echter minder positief voor de bewoners met ernstige cognitieve achteruitgang, bij wie de scores op deze zelfrapportage vragenlijsten minder samenhang vertonen met observaties van verdrietig gedrag, pijn, en sociale betrokkenheid dan bij bewoners met lichte cognitieve achteruitgang. De onderzoeker concludeert dat deze verschillen kunnen betekenen dat bewoners met ernstige cognitieve problematiek minder goed in staat zijn zelfrapportage vragenlijsten te beantwoorden of dat observatoren meer moeite hebben verdrietig gedrag, pijn en sociale betrokkenheid te observeren in deze groep. Zou het echter ook mogelijk kunnen zijn dat de geselecteerde vergelijkingsschalen minder relevant zijn voor mensen met ernstige cognitieve achteruitgang en derhalve resulteren in een andere score verdeling, die de correlatie beïnvloedt?

Al met al, een gedegen proefschrift dat niet alleen een interessante theorie biedt voor kwaliteit van leven in het verpleeghuis, maar ook een goede eerste aanzet vormt tot een instrumentarium voor het evalueren van kwaliteit van leven als uitkomst maat in de verpleeghuiszorg.

Dr. Saskia Teunisse

Hillegom, januari 2005

Naschrift

Het is een zeer leuke ervaring om een uitgebreide samenvatting van mijn proefschrift te lezen die is geschreven door een ander! Als deze dan ook nog gepaard gaat met een positieve beoordeling van de kwaliteit van het proefschrift kan ik niet anders dan zeer content zijn met deze

bespreking. Ik dank Saskia Teunisse hartelijk voor haar schrijven en ga met genoegen in op haar opmerkingen.

Haar kritische noot dat bij validatie van sommige schalen vergelijkingsschalen zijn gebruikt die in andere hoofdstukken niet optimaal valide bleken, wijst inderdaad op een tekortkoming die geldt voor het hele onderzoek. In dit soort onderzoek is te prefereren om de te beoordelen schalen te vergelijken met een gouden standaard. Een eerste probleem daarbij is echter dat gouden standaarden tijdsintensief zijn om toe te passen. Verder zijn voor constructen als 'sociale betrokkenheid' geen gouden standaarden bekend, terwijl voor 'depressie' wel gouden standaarden bestaan, maar daarbij bestaan twijfels over de toepasbaarheid in de verpleeghuispopulatie. Daarom zijn in deze studie zogenaamde 'vergelijkings-schalen' gebruikt waarvan de validiteit soms onbekend was. De schalen voor subjectief welbevinden bleken niet optimaal valide, met name bij bewoners met ernstige cognitieve stoornissen. Deze waren vaak wel in staat om de vragen te beantwoorden, terwijl de gegeven antwoorden op de verschillende schalen niet goed overeenkwamen. De situatie werd nog complexer toen bleek dat ook de observatieschalen voor subjectief welbevinden minder valide zijn in die groep, een bevinding waar ik zo op terugkom.

Een andere opmerking betreft de gevonden lage betrouwbaarheid van de MDS-depressieschaal (DRS). De DRS, één van de bovengenoemde schalen voor subjectief welbevinden, bleek in het onderzoek minder betrouwbaar en valide dan de MDS-schalen voor cognitie (CPS) en ADL (ADL-H). Die bevinding zal deels kunnen worden toegeschreven aan het verschijnsel dat psychologische aspecten van gedrag moeilijker te observeren zijn dan bijvoorbeeld ADL-functioneren. Saskia Teunisse vraagt zich af of de lagere betrouwbaarheid van de DRS een gevolg kan zijn van het grote tijdsinterval tussen de 2 metingen waarmee de betrouwbaarheid werd bepaald, wat zou kunnen duiden op goede responsiviteit. Dit is helaas niet het geval. Ik heb voor een subgroep waarbij het tijdsinterval 1 tot 2 weken was dezelfde resultaten gezien, iets dat geldt voor alle schalen. De betrouwbaarheid van de DRS was bovendien ook

duidelijk lager bij bewoners van een psychogeriatrische afdeling. Recent onderzoek in de VS zet ook vraagtekens bij de betrouwbaarheid en validiteit van deze DRS, en momenteel wordt dan ook door een internationale groep onderzoekers gewerkt aan een herziene versie.

De schalen voor subjectief welbevinden krijgen alle aandacht in hoofdstuk 7. Daar wordt gevonden dat de correlatie tussen zes schalen voor algemeen subjectief welbevinden lager zijn bij bewoners met ernstige dementie. Dit geldt voor zowel de observatieschalen als de zelfbeoordelingsschalen. Saskia Teunisse vraagt zich af of de beoordeelde schalen minder relevant zijn voor mensen met ernstige cognitieve achteruitgang. Dit is zeker mogelijk. Het meten van kwaliteit van leven (of psychisch welbevinden) bij mensen met cognitieve achteruitgang is een complex onderzoeksgebied, en de vraag wat psychisch welbevinden precies omvat bij mensen met ernstige dementie is nog niet beantwoord. Zoals in hoofdstuk 2 van het proefschrift uitgebreid is beschreven, wil ik er niet van uit gaan dat de inhoud van kwaliteit van leven fundamenteel verschilt in verschillende groepen mensen, maar het relatieve belang van domeinen kan wel verschillen. Desalniettemin denk ik dat de gevonden verschillen tussen de zes schalen vooral in de betrouwbaarheid van deze schalen moeten worden gezocht en minder in de constructvaliditeit ervan.

Tot slot wil ik graag een lans breken voor psychometrisch onderzoek in de ouderenzorg en de verpleeghuiszorg in het bijzonder. Veel schalen die worden gebruikt in de dagelijkse verpleeghuiszorg zijn niet gevalideerd voor deze populatie, zelfs een Geriatrische Depressie Schaal wordt nu pas, in het AGED-onderzoek van de Vrije Universiteit Amsterdam en de Rijksuniversiteit Groningen, gevalideerd voor de Nederlandse verpleeghuispopulatie. Het meten van kwaliteit van leven is momenteel erg populair, maar heel moeilijk. Voordat valide en betrouwbare schalen bestaan voor ouderen met verschillende capaciteiten is nog veel schaalonderzoek nodig. Ik nodig de lezer dan ook van harte uit hieraan bij te dragen.

Dr. Debby Gerritsen

Nijmegen, januari 2005