

Routine outcome monitoring bij de behandeling van angststoornissen: diagnosespecifieke versus generieke meetinstrumenten

M. VAN DER MHEEN, L.M. TER MORS, M.A. VAN DEN HOUT, D.C. CATH

- ACHTERGROND** Ggz-instellingen maken in toenemende mate gebruik van routine outcome monitoring (ROM) om de effectiviteit van behandeling vast te stellen. Het is onduidelijk welke instrumenten meer verandering registreren bij de toepassing van ROM bij angststoornissen: diagnosespecifieke of generieke meetinstrumenten.
- DOEL** Vaststellen van de mate waarin diagnosespecifieke versus generieke ROM-instrumenten vooruitgang registreren bij patiënten met een angststoornis.
- METHODE** Bij het Altrecht Academisch Angstcentrum werden 160 volwassenen met een angststoornis behandeld met cognitieve gedragstherapie. Ongeveer de helft ontving tevens medicatie. De patiënten voltooiden voor en na behandeling een meting met twee generieke instrumenten (Outcome Questionnaire (OQ-45) en Brief Symptom Inventory (BSI)) en twee diagnosespecifieke instrumenten afgestemd op de hoofddiagnose.
- RESULTATEN** De generieke BSI en diagnosespecifieke meetinstrumenten lieten in het algemeen een groter verschil tussen voor- en nameting zien dan de generieke OQ-45.
- CONCLUSIE** De BSI en diagnosespecifieke meetinstrumenten laten na cognitieve gedragstherapie sterkere verbeteringen zien dan de OQ-45. Mogelijk is de BSI bij angststoornissen een relatief specifieke maat. Bij het kiezen van een meetinstrument wordt aangeraden behandeldoelen (symptoomreductie of verbetering van kwaliteit van leven) en andere redenen van afname in beschouwing te nemen; als het doel is om de verandering van specifieke angstklachten in kaart te brengen, zijn BSI of specifieke maten te prefereren.

TIJDSCHRIFT VOOR PSYCHIATRIE 60(2018)1, 11-19

TREFWOORDEN angststoornis, diagnosespecifiek meetinstrument, generiek meetinstrument, routine outcome monitoring



Om de effectiviteit van behandelingen te meten, maken ggz-instellingen in toenemende mate gebruik van *routine outcome monitoring* (ROM; Nugter & Buwalda 2012). ROM houdt in dat behandelvooruitgang routinematig wordt vastgesteld met meetinstrumenten. Het gebruik van ROM is toegenomen vanwege de vraag van zorgverzekeraars naar meer inzichtelijkheid in de kwaliteit van behandelingen

(Hoenders e.a. 2013). Bovendien is de interesse in ROM gegroeid omdat onderzoek suggereert dat het routinematig meten van behandelvooruitgang en daarover feedback geven aan de patiënt de behandeluitkomst kan verbeteren (Miller e.a. 2005; Lambert 2007). Daarnaast is de wetenschappelijke interesse in ROM toegenomen door de overweging dat gerandomiseerde gecontroleerde trials (RCT's)

bepaalde externe validiteit hebben en dat kennis over het naturalistische beloop van behandelingen evenzeer waarde heeft (Gilbody e.a. 2002; zie Shafran e.a. 2009 en Lambert 2013 voor tegenargumenten).

Bij de toepassing van ROM kan gebruikgemaakt worden van meetinstrumenten die in meer of mindere mate diagnosespecifiek of generiek zijn. Diagnosespecifieke maten meten de aanwezigheid en ernst van specifieke psychiatrische symptomen (Luteijn e.a. 2011) en kunnen derhalve worden toegepast om behandelvoortgang en -uitkomst te meten in diagnostisch homogene groepen (Nugter & Buwalda 2012). Generieke maten meten daarentegen niet-diagnosespecifieke psychiatrische symptomen en kwaliteit van leven (Patrick & Deyo 1989) en maken vergelijkingen tussen en over verschillende diagnostische groepen heen mogelijk (Nugter & Buwalda 2012).

Het onderscheid tussen diagnosespecifieke en generieke maten is geen strikte dichotomie. Meetinstrumenten verschillen in mate van specificiteit, met bijvoorbeeld de *Outcome Questionnaire-45* (Lambert e.a. 1996) als generieke maat en de *Obsessive-Compulsive Inventory-Revised* (Foa e.a. 2002) als maat die specifiek betrekking heeft op obsessieve-compulsieve symptomen. Omdat generieke instrumenten, zoals gezegd, vergelijking tussen diagnostische groepen mogelijk maken en omdat het praktischer is om bij alle patiënten hetzelfde ROM-instrument te gebruiken, kan er een voorkeur ontstaan voor gebruik van generieke maten. De vraag is echter of generieke meetinstrumenten wel evenveel therapeutische voortgang registreren als diagnosespecifieke maten.

Om deze vraag te beantwoorden, verrichtten wij onderzoek waarbij de geïncludeerde patiënten allen werden behandeld met cognitieve gedragstherapie (GGT; naast antidepressiva een behandeling van eerste voorkeur bij angststoornissen, zie Trimbos-instituut 2013). Aangezien CGT bij angststoornissen zich primair richt op reductie van angstklachten en minder op algemene psychiatrische symptomen en algeheel psychosociaal welzijn, verwachten wij dat de behandeluitkomsten, dat wil zeggen de verschillen tussen voor- en nameting, groter zouden zijn op diagnosespecifieke maten dan op generieke maten.

METHODE

Participanten

Metingen werden verzameld van 172 patiënten met angststoornissen van dusdanige ernst en/of complexiteit dat zij waren verwezen naar een tweede- of derdelijns angstcentrum. Uit de oorspronkelijke groep werden 12 (6,9%) patiënten verwijderd wegens incomplete metingen (83,3% vrouw, 16,7% man; gemiddelde leeftijd 29,5; SD 8,6). Van deze patiënten hadden 8 (66,7%) als hoofddiagnose

AUTEURS

MALINDI VAN DER MHEEN, onderzoekspsycholoog, Researchafd. Kinder- en Jeugdpsychiatrie en Psychologie, Erasmus MC-Sophia kinderziekenhuis, Rotterdam.

LARA M. TER MORS, psycholoog.

MARCEL A. VAN DEN HOUT, klinisch psycholoog, Altrecht Academisch Angstcentrum en hoogleraar Klinische psychologie, Universiteit Utrecht.

DANIËLLE C. CATH, indertijd psychiater, hoofd Altrecht Academisch Angstcentrum; thans: universitair hoofddocent vakgroep Psychiatrie, Rob Giel Onderzoekscentrum Groningen en Universiteit Groningen; opleider psychiatrie GGZ Drenthe.

CORRESPONDENTIEADRES

Malindi van der Mheen, Erasmus MC-Sophia kinderziekenhuis, Researchafd. Kinder- en Jeugdpsychiatrie en Psychologie, KP-2841, Wytemaweg 8, 3015 CN Rotterdam.
E-mail: m.vandermheen@erasmusmc.nl

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 3-7-2017.

een obsessieve-compulsieve stoornis (OCS), 2 (16,7%) sociale angststoornis (SAS), 1 (8,3%) posttraumatische stressstoornis (PTSS) en 1 (8,3%) paniekstoornis (PS).

De uiteindelijke groep bestond uit 160 patiënten van 18 tot 59 jaar (gem. 34,9; SD 10,0), 97 (60,6%) vrouwen en 63 (39,4%) mannen. Diagnoses werden vastgesteld met de Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; APA 2000) en met het Gestructureerd Klinisch Interview voor DSM-IV-as I-stoornissen (SCID-I; First e.a. 1996). 44 (27,5%) patiënten hadden de hoofddiagnose OCS, 16 (10,0%) PTSS, 56 (35,0%) paniekstoornis met of zonder agorafobie (PS(A)) en 44 (27,5%) SAS. Bij 131 patiënten (81%) was er sprake van minstens één comorbide DSM-IV-diagnose en 80 (50,0%) patiënten hadden ten minste één comorbide angststoornis.

Behandeling

Het Altrecht Academisch Angstcentrum (AAA) behandelde de patiënten ambulantly middels CGT, eventueel aangevuld met *eye movement desensitisation and reprocessing* (EMDR). 81 (50,6%) patiënten kregen hiernaast farmacotherapie. Het aantal behandelingsessies varieerde van 3 tot 65 (gem. 19,3; SD 10,6).

Procedure

De data werden tussen 2008 en 2013 bij patiënten van het AAA verzameld. Patiënten vulden standaard voorafgaand

aan behandeling vragenlijsten in en op verzoek van behandelbaar ook na afloop van behandeling. Dit werd gedaan via internet (thuis of op de polikliniek). De meting bestond onder meer uit twee generieke meetinstrumenten en twee diagnosespecifieke meetinstrumenten afgestemd op hoofddiagnose.

Instrumenten

GENERIEKE MATEN

Outcome Questionnaire (OQ-45; Lambert e.a. 1996; de Jong e.a. 2008). Deze vragenlijst met 45 items meet angst en somatisch leed (d.w.z. distress), leed door symptomen, kwaliteit van interpersoonlijke relaties en sociale rol. Items worden beoordeeld op een 5-puntslikertschaal (0 = nooit, 4 = bijna altijd). De interne betrouwbaarheid is hoog ($\alpha = 0,96$) en de validiteit is voldoende (de Jong e.a. 2008). De OQ-45 is gevalideerd aan de *Symptom Checklist* (SCL-90; Arrindell & Ettema 1986), *Groningse Vragenlijst Sociaal Gedrag* (GVSG-45; de Jong & van der Lubbe 2001) en *Depression Anxiety Stress Scales* (DASS; de Beurs e.a. 2001).

Brief Symptom Inventory (BSI; Derogatis 1975; de Beurs & Zitman 2006). Deze vragenlijst (53 items) meet negen dimensies van psychiatrische symptomen: somatische klachten, cognitieve problemen, interpersoonlijke gevoeligheid, depressieve stemming, angst, hostiliteit, fobische angst, paranoïde gedachten en psychoticisme. Items worden beoordeeld op een 5-puntslikertschaal (0 = helemaal niet, 4 = heel erg). De interne betrouwbaarheid is hoog ($\alpha = 0,96$) en de validiteit is voldoende (de Beurs & Zitman 2006). De BSI is gevalideerd aan de SCL-90 (Arrindell & Ettema 1986) en negen diagnosespecifieke meetinstrumenten.

DIAGNOSESPECIFIEKE MATEN

OBSESSIEVE-COMPULSIEVE STOORNIS

Yale-Brown Obsessive Compulsive Scale (Y-BOCS; Goodman e.a. 1989a; Arrindell e.a. 1990). Deze vragenlijst meet obsessieve-compulsieve symptomen (10 items). Items worden beoordeeld op een 5-puntslikertschaal (0 = geen symptomen, 4 = extreme symptomen). De psychometrische kwaliteit van de Nederlandse vertaling is niet onderzocht. De oorspronkelijke versie is echter voldoende valide en betrouwbaar (Goodman e.a. 1989a; Goodman e.a. 1989b; Steketee e.a. 1996). De Nederlandse versie blijkt bruikbaar voor het meten van therapie-effect (Van Oppen e.a. 1995). *Obsessive-Compulsive Inventory-Revised* (OCI-R; Foa e.a. 2002; Cordova-Middelbrink e.a. 2007). Deze vragenlijst meet obsessieve-compulsieve symptomen (18 items). De zes subschalen zijn: wassen, controleren, ordenen, obsessies, verzamelen en neutraliseren. Items worden beoordeeld op een 5-puntslikertschaal (0 = helemaal niet, 4 = extreem). De

psychometrische kwaliteit van de Engelse versie is goed (Foa e.a. 2002). De psychometrische kwaliteit van de Nederlandse versie is niet onderzocht.

POSTTRAUMATISCHE STRESSSTOORNIS

PTSD Symptom Scale – Self-Report Version (PSS-SR; Arntz 1993; Foa e.a. 1993). Deze vragenlijst meet PTSD-symptomen (17 items). De drie subschalen zijn: herbeleving, vermijding en arousal. Items worden beoordeeld op een 4-puntslikertschaal (0 = helemaal niet, 3 = bijna altijd/meer dan vijf keer per week). De interne betrouwbaarheid is hoog ($\alpha = 0,93$). De validiteit is goed (Wohlfarth e.a. 2003).

Posttraumatic Cognitions Inventory (PTCI; Foa e.a. 1999; van Emmerik e.a. 2006). Deze vragenlijst meet traumagerelateerde gedachten en overtuigingen (36 items). De drie subschalen zijn: negatieve cognities over zichzelf, negatieve cognities over de wereld en zelfverwijt. Items worden beoordeeld op een 7-puntslikertschaal (1 = helemaal niet mee eens, 7 = helemaal mee eens). De interne betrouwbaarheid is hoog ($\alpha = 0,94$). De validiteit is goed (van Emmerik e.a. 2006).

PANIEKSTOORNIS MET OF ZONDER AGORAFOBIE

Mobility Inventory (MI; Chambless e.a. 1985; de Beurs 2001). Deze vragenlijst bestaat uit 26 items. De items hebben betrekking op situaties die vaak vermeden worden door patiënten met paniekstoornis en/of agorafobie. Mate van vermijding (alleen of met iemand samen) wordt beoordeeld op een 5-puntslikertschaal (0 = nooit, 4 = altijd). De psychometrische kwaliteit van de Nederlandse vertaling is niet onderzocht. De oorspronkelijke versie is echter voldoende valide en betrouwbaar (Chambless e.a. 1985).

Body Sensation Questionnaire (BSQ; Chambless e.a. 1984; Bouman 1998). Deze vragenlijst meet angst voor fysieke symptomen die gepaard kunnen gaan met angst (17 items). Vanwege gebrek aan een definitieve factoroplossing wordt alleen de totaalscore berekend. Items worden beoordeeld op een 5-puntslikertschaal (0 = niet angstig of bezorgd, 4 = heel erg angstig). De interne betrouwbaarheid is hoog ($\alpha = 0,91$). De validiteit is goed (Bouman 1998).

SOCIALEANGSTSTOORNIS

Social Phobia and Anxiety Inventory (SPAI; Turner e.a. 1989; Scholing e.a. 1995). Deze vragenlijst meet symptomen van sociale angst (45 items). De twee subschalen zijn: sociale fobie en agorafobie. Items worden beoordeeld op een 7-puntslikertschaal (1 = nooit, 7 = altijd). De betrouwbaarheid is hoog ($\alpha = 0,94$). De validiteit is goed (Scholing e.a. 1995).

Self-Consciousness Scale (SCS; Fenigstein e.a. 1975; Vleeming & Engelse 1981). Deze vragenlijst bestaat uit drie subschalen (23 items): persoonlijk zelfbewustzijn, publiek zelf-

bewustzijn en sociale angst. Items worden beoordeeld op een 5-puntslikertschaal (0 = heel karakteristiek/helemaal mee eens, 4 = helemaal niet karakteristiek/helemaal niet mee eens). De psychometrische kwaliteit van de Nederlandse versie is niet onderzocht. De oorspronkelijke versie is echter voldoende valide en betrouwbaar (Fenigstein e.a. 1975).

Statistische analyses

Van alle patiënten werden, bij zowel voormeting als name-ting, totaalscores vastgesteld. Alle patiënten scoorden op de OQ-45 en BSI. Patiënten met OCS scoorden daarnaast de Y-BOCS en OCI-R, patiënten met PTSS scoorden de PSS-SR en PTCI, patiënten met PS(A) scoorden de MI en BSQ en patiënten met SAS scoorden de SPAI en SCS.

Om de totaalscores op de verschillende meetinstrumenten te kunnen vergelijken, werden de scores gestandaardiseerd naar z-scores. Een z-score is een waarde uitgedrukt in het aantal standaarddeviaties dat het afwijkt van het gemiddelde (Field 2009). Voor ieder meetinstrument bedraagt het gemiddelde van de gestandaardiseerde voor- en name-ting tezamen 0 met een standaarddeviatie van 1. Om de statistische power te vergroten, werd één variabele gemaakt voor alle diagnosespecifieke maten door per patiënt de gestandaardiseerde totaalscores op de diagno-sespecifieke meetinstrumenten te middelen. Hierna bere-kenen wij verschilcores op de OQ-45, BSI en de diagno-sespecifieke maten door per patiënt de gestandaardiseerde nametingscore af te trekken van de gestandaardiseerde voormetingscore. Deze gemiddelde verschilcores geven een indicatie van behandeluitkomst.

De analyses werden in drie stappen verricht.

1. Allereerst werden de gestandaardiseerde verschilcores op de OQ-45, BSI en samengevoegde diagnosespecifieke

meetinstrumenten vergeleken met variantieanalyse (éénweg-herhaaldemetingen-ANOVA), om te onderzoe-ken of de mate waarin vooruitgang werd geregistreerd in het algemeen verschilde voor generieke versus diag-nosespecifieke maten.

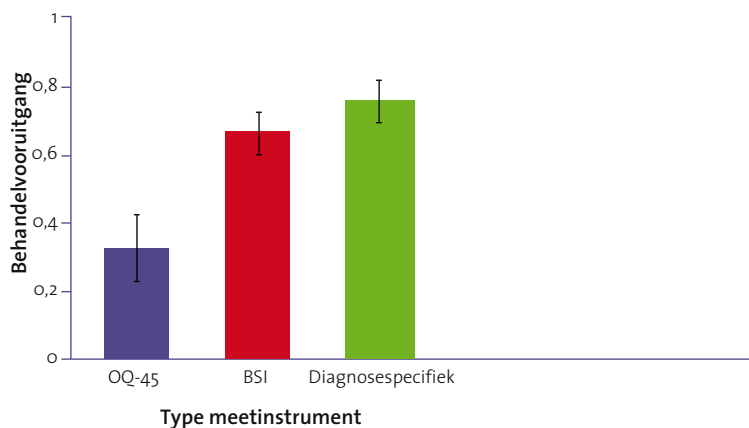
2. Om de vraag te beantwoorden of dit verschilde per hoofddiagnose en of behandeluitkomst ongeacht type meetinstrument per diagnose verschilde, werd in een tweede stap een 3x4 herhaaldemetingen-ANOVA verricht met meetinstrument (OQ-45, BSI of diagnosespecifieke maat) als binnengroepfactor en hoofddiagnose (OCD, PTSS, PS(A) of SAS) als tussengroepfactor.
3. Ten slotte werden de gestandaardiseerde verschilcores van de generieke en alle individuele diagnosespecifieke meetinstrumenten per hoofddiagnose vergeleken met behulp van vier éénweg-herhaaldemetingen-ANOVA's. Deze laatste analyse beantwoordde de vraag of de mate waarin alle individuele meetinstrumenten vooruitgang registreren, verschilde per hoofddiagnose.

RESULTATEN

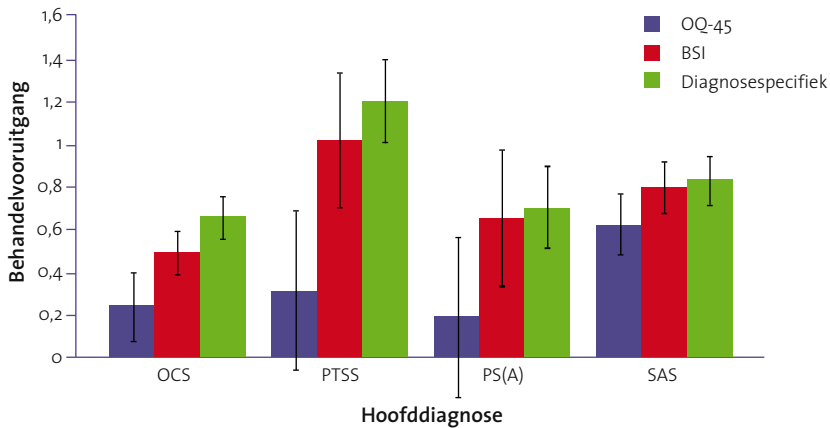
Generieke en samengevoegde specifieke instru-menten; ongeacht diagnose

In **FIGUUR 1** is de behandeluitkomst weergegeven, vastgesteld door de verschillende meetinstrumenten. Mauchly's test toonde dat de assumptie van sfericiteit geschonden was ($\chi^2(2) = 52,06$; $p < 0,001$). Aangezien de schatting van sferi-citeit hoger dan 0,75 was ($\epsilon = 0,79$), werden de vrijheidsgra-den aangepast met correctie volgens Huynh-Feldt (Field 2009). De resultaten van de éénweg-herhaaldemetingen-ANOVA lieten een significant verschil in registratie van ver-andering zien tussen de meetinstrumenten ($F(1,57, 250,32) = 9,96$; $p < 0,001$), met een matige effectgrootte (Pallant

FIGUUR 1 Behandeluitkomst weergegeven in gestandaardiseerde verschilcores gemeten met generieke meetinstrumenten (OQ-45 en BSI) en samengevoegde diagnosespecifieke meetinstrumenten; de foutenbalken geven de standaardfouten weer



FIGUUR 2 Behandeluitkomst weergegeven in gestandaardiseerde verschilscores gemeten met generieke meetinstrumenten (OQ-45 en BSI) en samengevoegde diagnosespecifieke meetinstrumenten per hoofddiagnose; de foutenbalken geven de standaardfouten weer



2010): $\eta_p^2 = 0,06$. Een bonferroni-post-hoc-toets toonde dat de BSI (gem. 0,67; SD 0,06) significant meer vooruitgang registreerde dan de OQ-45 (gem. 0,32; SD 0,10); $p = 0,01$. Ook bleken de diagnosespecifieke maten (gem. 0,76; SD 0,06) significant meer vooruitgang te registreren dan de OQ-45 ($p < 0,001$). De diagnosespecifieke maten bleken niet significant meer vooruitgang te registreren dan de BSI, $p = 0,61$.

Generieke vs. samengevoegde specifieke instrumenten; per diagnose

In **FIGUUR 2** is, per stoornis, de behandeluitkomst weergegeven, zoals vastgesteld met de OQ-45, BSI en samengevoegde diagnosespecifieke meetinstrumenten. Mauchly's test toonde dat de assumptie van sfericiteit geschonden was ($\chi^2(2) = 50,30$; $p < 0,001$). Daarom pasten we de vrijheidsgraden aan met correctie volgens Huynh-Feldt ($\epsilon = 0,81$). De resultaten van de 3x4 herhaaldemetingen-ANOVA (3 types meetinstrument x 4 hoofddiagnoses) toonden wederom een hoofdeffect voor type vragenlijst: $F(1,61, 251,09) = 10,35$; $p < 0,001$.

De verandering op de OQ-45 was het kleinst en die op de diagnosespecifieke maten het grootst. Ook was er een hoofdeffect voor hoofddiagnose ($F(3, 156) = 3,06$; $p = 0,03$), wat betekent dat er verschil was in behandeluitkomst per hoofddiagnose, ongeacht type meetinstrument. De effectgrootte was matig: $\eta_p^2 = 0,06$.

Een *least significant differences* (LSD)-post-hoc-toets wees uit dat patiënten met OCS (gem. 0,45; SD 0,51) een significant kleinere verandering tussen voor- en nameting lieten zien dan patiënten met PTSS (gem. 0,83; SD 0,80; $p = 0,03$) en dan patiënten met SAS (gem. 0,73; SD 0,54; $p = 0,02$). De overige diagnostische groepen verschilden niet significant in behandeluitkomst.

Er bleek geen significant interactie-effect te zijn tussen type meetinstrument en hoofddiagnose ($F(4,83, 251,09) = 0,69$; $p = 0,63$). Dit betekent dat het verschil in registratie van vooruitgang door de meetinstrumenten niet verschilden in de verschillende diagnostische groepen.

Generieke vs. afzonderlijke specifieke instrumenten; per diagnose

In **FIGUUR 3** is de behandeluitkomst gemeten door alle afzonderlijke meetinstrumenten per hoofddiagnose weergegeven.

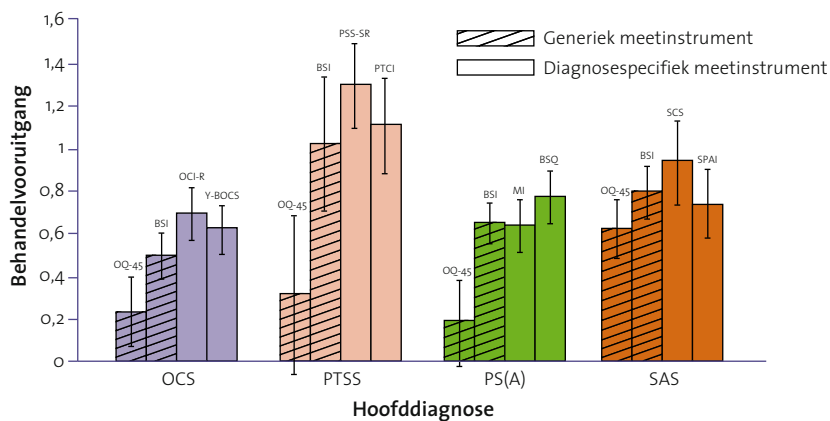
OBSESSIEVE-COMPULSIEVE STOORNIS

Voor de diagnose OCS ($n = 44$) toonde Mauchly's test dat de assumptie van sfericiteit geschonden was ($\chi^2(5) = 37,71$; $p < 0,001$). Aangezien de schatting van sfericiteit lager dan 0,75 was ($\epsilon = 0,64$), pasten we de vrijheidsgraden aan met correctie volgens Greenhouse-Geisser (Field 2009). Een éénweg-herhaaldemetingen-ANOVA toonde een niet-significant verschil in verandering tussen voor- en nameting bij patiënten met OCS, gemeten met de OQ-45, BSI, OCI-R en Y-BOCS ($F(1,91, 82,27) = 2,78$; $p = 0,07$). Het patroon suggereerde echter dat, wederom, de verandering op diagnosespecifieke maten groter was dan op generieke maten.

POSTTRAUMATISCHE STRESSSTOORNIS

Voor de diagnose PTSS ($n = 16$) toonde Mauchly's test dat de assumptie van sfericiteit geschonden was ($\chi^2(5) = 24,60$; $p < 0,001$). Daarom werden de vrijheidsgraden aangepast met correctie volgens Greenhouse-Geisser ($\epsilon = 0,48$). Een éénweg-herhaaldemetingen-ANOVA toonde een niet-significant verschil in registratie van vooruitgang ($F(1,44, 21,66) = 2,94$; $p = 0,09$), waarbij de diagnosespecifieke PSS-SR de meeste vooruitgang leek te registreren.

FIGUUR 3 Behandeluitkomst weergegeven in gestandaardiseerde verschilscores per hoofddiagnose gemeten met generieke en afzonderlijke diagnosespecifieke meetinstrumenten; de foutenbalken geven de standaardfouten weer



PANIEKSTOORNIS MET OF ZONDER AGORAFOBIE

Voor de diagnose ps(A) ($n = 56$) toonde Mauchly's test dat de assumptie van sfericiteit geschonden was ($\chi^2(5) = 31,24$; $p < 0,001$). Daarom werden de vrijheidsgraden aangepast met correctie van Greenhouse-Geisser ($\epsilon = 0,74$). Een één-weg-herhaaldemetingen-ANOVA liet een significant verschil zien in behandeluitkomst van patiënten met ps(A), vastgesteld met verschillende meetinstrumenten ($F(2,22, 122,29) = 3,49$; $p = 0,03$). De effectgrootte was matig ($\eta_p^2 = 0,06$). Een LSD-post-hoc-toets toonde dat de BSQ (gem. 0,75; SD 0,92) significant meer vooruitgang registreerde dan de OO-45 (gem. 0,18; SD 1,5; $p = 0,02$). Ook de MI (gem. 0,62; SD 0,91) bleek significant meer vooruitgang te registreren dan de OO-45 ($p = 0,04$).

SOCIALEANGSTSTOORNIS

Voor de diagnose SAS ($n = 44$) toonde Mauchly's test dat de assumptie van sfericiteit geschonden was ($\chi^2(5) = 24,15$; $p < 0,001$). Daarom pasten we de vrijheidsgraden aan met correctie volgens Greenhouse-Geisser ($\epsilon = 0,74$). Er was geen significant verschil in behandeluitkomst van patiënten met SAS, gemeten met de OO-45, BSI, SCS en SPAI ($F(2,22, 95,51) = 0,73$; $p = 0,50$).

DISCUSSIE

Het doel van deze studie was om vast te stellen of diagnosespecifieke en generieke instrumenten die in ROM bij angststoornissen kunnen worden toegepast, verschillen in de mate waarin deze vooruitgang registreren. Patiënten vulden voor en na behandeling twee generieke meetinstrumenten (OO-45 en BSI) en twee diagnosespecifieke meetinstrumenten in. Door gestandaardiseerde verschilscores te vergelijken, verkregen we een indicatie van de mate waarin de meetinstrumenten vooruitgang registreren.

De opbouw van het artikel loopt van grove analyse naar verfijndere analyse (ZIE FIGUUR 1, 2 en 3). De gebruikte schalen bestaan, deels, uit subschalen die apart gescoord kunnen worden. Vooral met oog op de toegankelijkheid en overzichtelijkheid hebben we niet gekozen voor verdere detaillering door analyses van scores op subschalen. Wij verwachten dat dit voor het patroon in de data (hoe specifiek gericht op de klachten, hoe sterker de effecten) geen noemenswaardig verschil zal maken.

Verschillen in generieke instrumenten

De resultaten tonen een vrij duidelijk patroon. De (generieke) BSI en de samengevoegde diagnosespecifieke meetinstrumenten lieten grotere verschillen tussen voor- en nameting zien dan de (generieke) OO-45. De registratie van vooruitgang door de BSI en diagnosespecifieke instrumenten verschilde niet. Dit resultaat verschilde niet per hoofddiagnose.

Daarnaast hebben we de registratie van vooruitgang door de generieke en de afzonderlijke diagnosespecifieke meetinstrumenten met elkaar vergeleken per diagnostische groep. In deze analyse zijn alleen verschillen in registratie van vooruitgang gevonden tussen meetinstrumenten bij patiënten met ps(A): de diagnosespecifieke MI en BSQ bleken grotere verschillen tussen voor- en nameting te detecteren dan de generieke OO-45. Bij meetinstrumenten voor OCS en PTSS werden trends aangetroffen (zie FIGUUR 3). Interessant is dat de generieke meetinstrumenten (OO-45 en BSI) onderling verschillen in mate van registratie van vooruitgang wanneer alle diagnostische groepen samen worden geanalyseerd. Een verklaring hiervoor zou kunnen zijn dat de BSI op itemniveau meer gericht is op symptomen (bijv. 'bang zijn om te reizen met bus, trein of tram', 'opvliegers of koude rillingen'; de Beurs & Zitman 2006), terwijl de OO-45 meer items heeft die gericht zijn op

algeheel functioneren (bijv. 'ik ben een tevreden mens', 'ik heb het gevoel dat het niet goed gaat met mijn werk/schoolwerk'; de Jong e.a. 2008). Bovendien meten twee van de negen subschalen (11 van de 53 items) van de BSI (fobische) angst. De OQ-45 bevat daarentegen geen subschalen die enkel angst meten. De subschaal angst en somatisch leed van de OQ-45 bestaat uit algemenere items (bijv. 'ik heb het gevoel dat er iets ergs gaat gebeuren', 'ik lijd aan hoofdpijn') dan de subschalen angst en fobische angst van de BSI (bijv. 'bepaalde dingen, plaatsen of activiteiten vermijden omdat je er angstig van wordt', 'aanvallen van angst of paniek').

Aangezien CGT gericht is op reductie van angstklachten, is het niet verrassend dat de BSI bij angststoornissen meer verandering lijkt te detecteren dan de OQ-45. Dit verklaart mogelijk ook dat de mate van registratie van vooruitgang van de BSI vergelijkbaar is met de diagnosespecifieke meetinstrumenten. De figuren laten zien dat niet alleen in de groep als geheel (zie FIGUUR 1), maar ook in de vier subgroepen (zie FIGUUR 2 en 3) de gemeten vooruitgang groter is op de BSI dan op de OQ-45. De omstandigheid dat het verschil alleen significant was in de groep als geheel, is toe te schrijven aan grotere statistische power in de totale groep ($n = 160$) vergeleken met de subgroepen ($n = 16-56$). Een poweranalyse met behulp van G*Power (Faul e.a. 2007) suggereert dat iedere patiëntengroep moet bestaan uit ongeveer 50 personen om significante resultaten te behalen.

Verschillen per diagnosegroep

Bij de patiënten met een paniekstoornis was het verschil in registratie van vooruitgang tussen de instrumenten significant. Bij de patiënten met PTSS en OCS was er een (niet significante) trend in deze richting, maar bij de patiënten met sociale angst was er geen aanwijzing voor een verschil tussen de instrumenten ($p = 0,50$). Een verklaring kan zijn dat twee van de vier subschalen van de OQ-45 sociale rol en interpersoonlijke relaties meten, onderwerpen die raken aan de problematiek van patiënten met sociale angst. Voor deze groep geldt de OQ-45 wellicht als (relatief) diagnosespecifieke maat.

Ongeveer 80% van de patiënten voldeed aan de criteria van ten minste één comorbide stoornis. Het zou interessant zijn de ontwikkeling op de diagnosespecifieke maten van die comorbide stoornissen te documenteren. Daartoe zijn wij niet in staat, omdat diagnosespecifieke metingen alleen werden verricht voor de hoofddiagnose.

Implicaties voor ROM-praktijk

Wat zijn de implicaties van deze bevindingen voor de ROM-praktijk? De keuze van een evaluatie-instrument hangt af van het doel van de behandeling.

Als patiënt en behandelaar zich tot doel stellen psychische klachten, zoals samengevat in bijvoorbeeld een DSM-classificatie, te reduceren, dan zijn diagnosespecifieke maten te verkiezen. Bij patiënten met een angststoornis kan de BSI goede diensten bewijzen in de ROM-praktijk: het instrument is wat minder specifiek dan de diagnosespecifieke instrumenten, maar de registratie van vooruitgang is vergelijkbaar en de BSI laat vergelijking tussen verschillende patiëntengroepen toe.

Wanneer het behandelgoal is om ook algemene aspecten van kwaliteit van leven te bevorderen, kunnen generieke maten worden ingezet. Verbetering van algemene aspecten van kwaliteit van leven, zoals rollen in de samenleving, kan uiteraard ook de inzet van therapie zijn, maar is in sterke mate afhankelijk van de sociaaleconomische situatie van patiënt en van premorbide functioneren. Als die beide al minimaal waren, is een vooruitgang op de OQ-45 niet te verwachten, en lijkt deze maat als proxy voor de kwaliteit van psychotherapie binnen een instelling en voor vergelijking ervan tussen instellingen inadequaat en zelfs unfair.

Aantrekkelijk aan de OQ-45 en vergelijkbare meetinstrumenten is dat ze de mogelijkheid bieden behandel-effecten bij psychiatrische patiënten met uiteenlopende psychiatrische klachten te meten en te vergelijken. Maar wie, al dan niet als onderdeel van ROM, deze algemene 'quality of life'-metingen verricht, moet zich realiseren dat de effecten een forse onderschatting kunnen zijn van de vooruitgang die patiënten boeken in de therapie.

LITERATUUR

- American Psychiatric Association. Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington: APA; 2000.
- Arntz A. Dutch translation of the PSS-SR. Maastricht: Universiteit Maastricht; 1993.
- Arrindell WA, Albersnagel FA, Oppen P van. Yale-Brown Obsessive Compulsive Scale (Y-BOCS). Ned. bewerking, verkorte versie. 1990. http://www.tijdschriftvoordepsychiatrie.nl/assets/measuringinstruments/meetinstrumenten_130pdf.pdf.
- Arrindell WA, Ettema JHM. SCL-90: Handleiding bij een multidimensionele psychopathologie-indicator. Lisse: Swets & Zeitlinger; 1986.
- Beurs E de. Mobility Inventory (Ned. vert.). Leiden: Leids Universitair Medisch Centrum; 2001.
- Beurs E de, Dyck R van, Marquenie LA, Lange A, Blonk RWB. De DASS: een vragenlijst voor het meten van depressie, angst en stress. Gedragstherapie 2001; 34: 35-53.
- Beurs E de, Zitman FG. De Brief Symptom Inventory (BSI): de betrouwbaarheid en

- validiteit van een handzaam alternatief voor de sci-90. *Maandbl Geest Volksgezondh* 2006; 61: 120-41.
- Bisson JI, Ehlers A, Matthews R, Pilling S, Richards D, Turner S. Psychological treatments for chronic post-traumatic stress disorder: Systematic review and meta-analysis. *Br J Psychiatry* 2007; 190: 97-104.
 - Bouman TK. Kort instrumenteel: de Body Sensation Questionnaire (BSQ). *Gedragstherapie* 1998; 31: 165-8.
 - Chambless DL, Caputo GC, Bright P, Gallagher R. Assessment of fear of fear in agoraphobics: the body sensations questionnaire and the agoraphobic cognitions questionnaire. *J Consult Clin Psychol* 1984; 52: 1090-7.
 - Chambless DL, Caputo GC, Jasin SE, Gracely EJ, Williams C. The mobility inventory for agoraphobia. *Behav Res Ther* 1985; 23: 35-44.
 - Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale: Erlbaum; 1988.
 - Cordova-Middelbrink JA, Dek ECP, Engelbarts MMB. *Automatisering bij subklinische obsessief-compulsieve stoornis* [Ongepubliceerde masterthesis]. Universiteit Utrecht; 2007.
 - Derogatis LR. *The Brief Symptom Inventory*. Baltimore: Clinical Psychometric Research; 1975.
 - Emmerik AAP van, Schoorl M, Emmelkamp PMG, Kamphuis JH. Psychometric evaluation of the Dutch version of the posttraumatic cognitions inventory (PTCI). *Behav Res Ther* 2006; 44: 1053-65.
 - Faul F, Erdfelder E, Lang AG, Buchner A. *G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences*. *Behav Res Methods* 2007; 39: 175-91.
 - Fenigstein A, Scheier MF, Buss AH. Public and private self-consciousness: Assessment and theory. *J Consult Clin Psychol* 1975; 43: 522-7.
 - Field A. *Discovering statistics using SPSS*. London: SAGE; 2009.
 - Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, e.a. The Obsessive-Compulsive Inventory: development and validation of a short version. *Psychol Assessment* 2002; 14: 485-95.
 - Foa EB, Riggs DS, Dancu CV, Rothbaum BO. Reliability and validity of a brief instrument for assessing post-traumatic stress disorder. *J Trauma Stress* 1993; 6: 459-73.
 - Foa EB, Tolin DF, Ehlers A, Clark DM, Orsillo SM. The posttraumatic cognitions inventory (PTCI): Development and validation. *Psychol Assessment* 1999; 11: 303-14.
 - Gilbody SM, House AO, Sheldon TA. Outcomes research in mental health: Systematic review. *Br J Psychiatry* 2002; 181: 8-16.
 - Goodman WK, Price LH, Rasmussen SA, Mazure C, Fleischmann RL, Hill CL, e.a. The Yale-Brown Obsessive Compulsive Scale. I. Development, use, and reliability. *Arch Gen Psychiatry* 1989a; 46: 1006-11.
 - Goodman WK, Price LH, Rasmussen SA, Mazure C, Delgado P, Heninger GR, e.a. The Yale-Brown Obsessive Compulsive Scale. II. Validity. *Arch Gen Psychiatry* 1989b; 46: 1012-6.
 - Hermann RC. Linking outcome measurement with process measurement for quality improvement. In: Ishak WW, Burt T, Sederer LI, red. *Outcome measurement in psychiatry, a critical review*. Washington: American Psychiatric Publishing; 2005. p. 23-34.
 - Hermans D, Eelen P, Orlemans H. *Inleiding tot de gedragstherapie*. Houten: Bohn Stafleu van Loghum; 2007.
 - Hoenders RHJ, Bos EH, Bartels-Velthuis AA, Vollbehn NK, Ploeg K van der, Jonge P de, e.a. Pittfalls in the assessment, analysis, and interpretation of routine outcome monitoring (ROM) data: results from an outpatient clinic for integrative mental health. *Adm Policy Ment Health* 2013; 41: 647-59.
 - Jong A de, Lubbe PM van der. *Groningse vragenlijst voor sociaal gedrag: zelfbeoordelingsvragenlijsten voor het vaststellen van problemen in het interpersoonlijk functioneren (handleiding)*. Groningen: Rob Giel Onderzoekscentrum; 2001.
 - Jong K de, Nugter MA, Polak MG, Wagenborg JEA, Spinhoven P, Heise WJ. De Nederlandse versie van de Outcome Questionnaire (OQ-45): een crossculturele validatie. *Psychol Gezondheid* 2008; 36: 35-45.
 - Laane R, Luijk R. ROM en de positie van zorgverzekeraars. *Tijdschr Psychiatr* 2012; 54: 135-9.
 - Lambert MJ. Presidential address: what we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychother Res* 2007; 17: 1-14.
 - Lambert MJ. *Bergin and Garfield's Handbook of psychotherapy and behavior change*. Hoboken: Wiley; 2013.
 - Lambert MJ, Burlingame GM, Umphress V, Hansen NB, Vermeersch DA, Clouse GC, e.a. The reliability and validity of the Outcome Questionnaire. *Clin Psychol Psychother* 1996; 3: 249-58.
 - Luteijn F, Barelids DPH, Arrindell WA, Deelman BG, Kamphuis JH, Vertommen H. *Psychologische diagnostiek in de gezondheidszorg*. Den Haag: Boom Lemma; 2011.
 - Miller SD, Duncan BL, Sorrel R, Brown GS. The partners for change outcome management system. *J Clin Psychol* 2005; 186: 199-208.
 - Noom MJ, De Jong K, Tiemens B, Kamsteeg F, Markus MT, Pot AM, e.a. Routine outcome monitoring en benchmarking: hoe kunnen we behandelresultaten op een zorgvuldige manier vergelijken? *Tijdschr Psychiatr* 2012; 54: 141-5.
 - Nugter MA, Buwalda VJA. Achtergronden en gebruiksmogelijkheden van ROM in de ggz. *Tijdschrift Psychiatr* 2012; 54: 111-8.
 - Oppen P van, Emmelkamp PMG, Balkom AJLM van, Dyck R van. The sensitivity to change of measures for obsessive-compulsive disorder. *J Anxiety Disord* 1995; 9: 241-8.
 - Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Med Care* 1989; 27: 217-32.
 - Pool G, Heuvel F, Ranchor AV, Sanderman R. *Handboek psychologische interventies bij chronisch-somatische aandoeningen*. Assen: Van Gorcum; 2004.

- Scholing A, Bögels SM, Van Velzen C. The Dutch Social Phobia and Anxiety Inventory (SPAI). Authorized translation; 1995.
- Shafran R, Clark DM, Fairburn CG, Arntz A, Barlow DH, Ehlers, e.a. Mind the gap: improving the dissemination of CBT. *Behav Res Ther* 2009; 47: 902-9.
- Trimbos-instituut. Multidisciplinaire richtlijn angststoornissen (3e revisie). Utrecht: Trimbos-instituut; 2013.
- Turner SM, Beide, DC, Dancu CV, Stanley MA. An empirically derived inventory to measure social fears and anxiety: the social phobia and anxiety inventory. *J Consult Clin Psychol* 1989; 1: 35-40.
- Vleeming RG, Engels JA. Assessment of private and public self-consciousness: a Dutch replication. *J Pers Assess* 1981; 45: 385-9.
- Walburg JA, Brinkmann J. Kwaliteit is geen toeval: de praktijk van de kwaliteitszorg in de ggz. Utrecht: GGZ Nederland; 2001.
- Wiger DE, Solberg KB. Tracking mental health outcomes: A therapist's guide to measuring client progress, analyzing data, and improving your practice. New York: Wiley; 2001.
- Wohlfarth TD, Brink W van den, Winkel FW, Smitten M ter. Screening for posttraumatic stress disorder: an evaluation of two self-report scales among crime victims. *Psychol Assess* 2003; 15: 101-9.

SUMMARY

Routine outcome monitoring in anxiety disorders: diagnosis-specific versus generic assessment instruments

M. VAN DER MHEEN, L.M. TER MORS, M.A. VAN DEN HOUT, D.C. CATH

BACKGROUND Mental health institutions increasingly utilise routine outcome monitoring (ROM) in order to assess treatment effectiveness. It remains unclear which instruments are more sensitive to change with regard to ROM in patients with an anxiety disorder: diagnosis-specific or generic instruments.

AIM To compare the sensitivity of diagnosis-specific versus generic ROM-instruments in patients with an anxiety disorder.

METHOD 160 adults with an anxiety disorder received cognitive behavioural therapy at the Altrecht Academic Anxiety Centre. Approximately half of the population also received medication. Patients completed an assessment both before and after treatment. This consisted of two generic instruments (Outcome Questionnaire (OQ-45) and Brief Symptom Inventory (BSI)) and two diagnosis-specific instruments, determined by the main diagnosis.

RESULTS The differences between pre- and post-treatment assessments were generally larger for the generic BSI and diagnosis-specific instruments than for the generic OQ-45.

CONCLUSION When assessed after cognitive behavioural therapy, the (generic) BSI and diagnosis-specific instruments indicated larger progress than the OQ-45. The BSI might be a relatively diagnosis-specific measure for anxiety disorders. When selecting an instrument for assessment, both the intended goal of treatment (symptom reduction or improvement in quality of life) and other reasons for assessment should be taken into consideration. The BSI or diagnosis-specific instruments are preferred if the goal is to assess the change in specific anxiety symptoms.

TIJDSCHRIFT VOOR PSYCHIATRIE 60(2018)1, 11-19

KEY WORDS anxiety disorder, diagnosis-specific instruments, generic instruments, routine outcome monitoring