

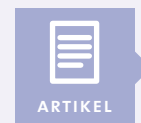
De PTSS Checklist voor DSM-5 (PCL-5) voor ROM: een vergelijking van de responsiviteit met de Outcome Questionnaire (OQ-45) en praktische bruikbaarheid

E. DE BEURS, K. THOMAES, H. KRONEMEIJER, J. DEKKER

- ACHTERGROND** De PTSS Checklist voor de DSM-5 (PCL-5) is mogelijk een goede aanvulling voor routine outcome monitoring (ROM) bij patiënten met PTSS.
- DOEL** Nagaan of de toepassing van de PCL-5 een nuttige aanvulling is die opweegt tegen de extra moeite die het invullen vergt van de patiënt.
- METHODE** Gegevens van de eerste en de laatste meting op de PCL-5 en de OQ-45 werden op individueel patiëntniveau vergeleken bij een groep van 464 patiënten van het Sinai Centrum van Arkin.
- RESULTATEN** De samenhang van scores op de instrumenten was hoog en variantieanalyse voor herhaalde metingen bracht geen verschil in responsiviteit aan het licht. Vergelijking van Cohens d (0,49 vs. 0,43) en Delta T (5,0 vs. 4,4) wees op een iets betere responsiviteit van de PCL-5 en ook de proportie herstelde patiënten was groter volgens de PCL-5 dan volgens de OQ-SD.
- CONCLUSIE** Op het eerste gezicht zijn de PCL-5 en de OQ-SD even gevoelig om verandering tijdens de behandeling te detecteren. De PCL-5 geeft echter meer gedetailleerde informatie over aard en ernst van de symptomen bij een individuele patiënt en met de PCL-5 zijn we iets beter in staat om herstel aan te tonen dan met de OQ-SD. We bevelen de toevoeging van de PCL-5 aan voor ROM bij patiënten met PTSS.

TIJDSCHRIFT VOOR PSYCHIATRIE 62(2020)6, 448-456

TREFWOORDEN OQ-45, PCL-5, posttraumatische stressstoornis, responsiviteit, ROM



ARTIKEL



Voor routine outcome monitoring (ROM; De Beurs e.a. 2011) ter ondersteuning van de behandeling en voor kwaliteitsmanagement zijn goede gegevens nodig over wat de behandeling oplevert (De Beurs e.a. 2017). Heeft mijn patiënt baat bij wat we doen en bespreken in de behandeling? Op welke aspecten van de symptomen boekt hij/zij vooral vooruitgang? Wat word ik wijzer als ik stilsta bij mijn caseload en succesvolle en minder succesvolle behan-

delingen overzie en evalueer? Wat vertellen geaggregeerde gegevens mij over de kwaliteit van de zorg die we leveren met mijn team of in mijn instituut? Zijn onze behandelresultaten beter dan die van vorig jaar?

Voor het beantwoorden van deze vragen kan ROM worden toegepast. Daar zijn meetinstrumenten voor nodig die bruikbare informatie bieden om de behandeling bij te sturen en die responsief genoeg zijn om verandering te

detecteren. Voor de selectie van meetinstrumenten is een belangrijke kwestie of stoornisspecifieke meetinstrumenten meerwaarde hebben ten opzichte van generieke instrumenten (Dingemans & Van Furth 2017; Van der Mheen e.a. 2018; Nugter e.a. 2019).

Het Sinai Centrum van Arkin gebruikte voor ROM aanvankelijk alleen de *Outcome Questionnaire* (OQ-45; Lambert e.a. 1996), maar sinds mei 2017 is de *Posttraumatische Stress Stoornis Checklist* voor de DSM-5 (PCL-5; Blevins e.a. 2015) aan het ROM-instrumentarium toegevoegd (Aarts e.a. 2019). De PCL-5 komt voort uit de PCL, het meest gebruikte zelfrapportage meetinstrument voor het meten van de aard en ernst van posttraumatische stressstoornis (Weathers e.a. 1993, 2013). De PCL is een zelfrapportage-instrument waarin alle symptomen van PTSS volgens de *Diagnostic en Statistical Manual of Mental Disorders* (DSM-5; American Psychiatric Association 2013) worden uitgevraagd. Met de introductie van de DSM-5 is de PCL gereviseerd en met 3 items uitgebreid tot de 20-item-PCL-5. De PCL-5 heeft goede psychometrische eigenschappen: een goede interne consistentie van de totale schaal (Cronbachs $\alpha = 0,94$), goede test-hertestbetrouwbaarheid ($r = 0,82$) en convergente en discriminante validiteit is aangetoond (Blevins e.a. 2015; Ashbaugh e.a. 2016). De Nederlandse vertaling van de PCL-5 is van Boeschoten e.a. (2014) van de Stichting Centrum '45. Zij repliceerden de goede psychometrische eigenschappen van het oorspronkelijke instrument.

De PCL-5 past qua meetpretentie (ernst van de PTSS-symptomen) beter bij de doelgroep met PTSS dan de OQ-45, maar er is nog niet onderzocht of de PCL-5 ook geschikt is voor effectmeting en kwaliteitsmonitoring op basis van geaggregeerde gegevens. Naar verwachting leidt de PCL-5 tot een hogere Delta T dan de OQ, omdat dit stoornisspecifieke meetinstrument beter de veranderingen meet die in de behandeling worden nagestreefd. Een vergelijkbare situatie speelde bij patiënten met een eetstoornis. Dingemans en van Furth (2017) lieten zien dat de *Eating Disorders Questionnaire* (EDE-Q; Fairburn & Beglin 2008) aanzienlijk gevoeliger was voor therapie-effect dan de BSI (De Beurs & Zitman 2006).

De responsiviteit van instrumenten wordt het best vergeleken in een directe vergelijking: twee instrumenten worden dan simultaan herhaald afgenomen (De Beurs e.a. 2012). Indien de PCL-5 responsiever is dan de OQ-45, dan is de PCL-5 ook een beter uitgangspunt om de behandeluitkomst bij groepen patiënten te meten. Dat is belangrijk voor toepassing van de PCL-5 in gecontroleerd effectonderzoek, in observationeel onderzoek en voor kwaliteitsmanagement (Hermann 2007; Kilbourne e.a. 2018).

AUTEURS

EDWIN DE BEURS, hoogleraar, Sectie Klinische Psychologie, Universiteit Leiden, Leiden, en senior onderzoeker, afd. Zorg Onderzoek en Kwaliteit, GGZ Arkin, Amsterdam.

KATHLEEN THOMAES, psychiater, Sinai Centrum, GGZ Arkin, Amstelveen, onderzoeker, afd. Psychiatrie, Anatomie en Neuroscience, Amsterdams Universitair Medisch Centrum, locatie Vrije Universiteit, Amsterdam.

HANS KRONEMEIJER, onderzoeker, afd. Zorg Onderzoek en Kwaliteit, GGZ Arkin, Amsterdam.

JACK DEKKER, senior onderzoeker, afd. Zorg Onderzoek en Kwaliteit, GGZ Arkin, Amsterdam, en hoogleraar Klinische psychologie, Vrije Universiteit, Amsterdam.

CORRESPONDENTIEADRES

Edwin de Beurs, Sectie Klinische Psychologie, Universiteit Leiden, Wassenaarseweg 52, 2333 AK, Leiden.
E-mail: e.de.beurs@fsw.leidenuniv.nl

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 18-12-2019.

METHODE

Respondenten

Aan dit onderzoek namen 464 patiënten deel die behandeld werden voor PTSS bij het Sinai Centrum. De gemiddelde leeftijd van de groep was 44,6 (SD: 13,2); 52,4% van de patiënten was vrouw. Voor 440 patiënten (94,8% van de gehele groep) was diagnostische informatie beschikbaar; 306 (65,9%) hadden PTSS als primaire diagnose, 34 (7,3%) een persoonlijkheidsstoornis, 26 (5,6%) een angststoornis niet anders omschreven en 74 patiënten (15,9%) hadden een andere primaire diagnose.

Aan alle patiënten werd bij de intake gevraagd om toestemming voor gebruik van geanonimiseerde ROM-data voor wetenschappelijk onderzoek. De onderzochte groep vormde een selectie van alle patiënten die werden behandeld: er moesten een eerste meting en een herhaalmeting zijn met zowel de PCL-5 als de OQ-45. Bij 301 patiënten werden alleen de 25 items van de OQ-SD-subschaal aangeboden, door 163 patiënten werd de gehele OQ-45 ingevuld.

De PCL-5 en de OQ werden herhaaldelijk afgenomen met een tussenliggend interval dat verschilde tussen patiënten van 4 tot maximaal 99 weken. De gemiddelde lengte van het interval (inclusief een wachtlijstperiode voor veel patiënten) was 40,9 (SD: 23,2) weken. Wanneer er meerdere hermetingen waren, werd de laatste hermeting geselecteerd

waarbij beide instrumenten beschikbaar waren. Behandelingen bij Sinai duren meestal langer dan 4 weken; doorgaans 12 tot 20 weken en de hermetingen vonden soms tijdens en soms na afsluiting van de behandeling plaats. De gegevens geven dan ook geen juiste afspiegeling van het gehele behandelingseffect bij Sinai, maar kunnen wel dienen om de responsiviteit van de meetinstrumenten te vergelijken.

Instrumenten

De *PTSS Checklist voor DSM-5 (PCL-5)* bestaat uit 20 items die corresponderen met symptomen van PTSS volgens de DSM-5. Respondenten moeten per symptoom aangeven hoeveel last ze de afgelopen maand hadden op een schaal met vijf antwoordopties (0 = 'helemaal niet' tot 4 = 'extreem veel'). De totaalscore heeft een range van 0 tot 80. Er zijn vier subschalen die corresponderen met de symptoomclusters van de DSM-5: cluster B (herbeleving, 5 items: 1-5), cluster C (vermijding, 2 items: 6-7), cluster D (negatieve cognities en stemming, 7 items: 8-14) en cluster E (arousal, 6 items: 15-20).

De *Outcome Questionnaire (OQ-45)* meet de uitkomst van de behandeling met 45 items waarmee respondenten worden bevraagd over hoe vaak (op een schaal van 0 = 'nooit' tot 4 = 'bijna altijd') een symptoom voorkwam of optrad in de afgelopen week. 25 items gaan over symptomatische distress (OQ-SD), 11 meten functioneren in interpersoonlijke relaties (IR); 9 gaan over functioneren in de sociale rol (SR) op werk of opleiding. Ten slotte is er nog een subset van 13 SD-items die angst en somatische distress (ASD) meten. De OQ-45 is een betrouwbaar en valide meetinstrument met goede psychometrische eigenschappen (Lambert e.a. 1996) en deze zijn behouden in de Nederlandse vertaling (De Beurs e.a. 2005, De Jong e.a. 2007).

T-scoreconversie

De (sub)schalen van de PCL-5 en de OQ-45 worden berekend door de score op items op te tellen en hebben daardoor elk een andere schaal (metric) met een ander gemiddelde, spreiding en bereik van scores. Om scores vergelijkbaar te maken converteerden wij ze naar T-scores (De Beurs & Flens 2017). Dit zijn scores met een gemiddelde van 50 en een standaarddeviatie van 10 (McCall 1922). De T-scores zijn gekalibreerd op de gezonde populatie (De Beurs e.a. 2019a). Er zijn in Nederland nog geen gegevens van de PCL-5 verzameld in een bevolkingssteekproef; daarom gebruikten we Canadese gegevens over gemiddelde en spreiding (Ashbaugh e.a. 2016). Voor de OQ gebruikten we normgegevens voor de Nederlandse bevolking (Timman e.a. 2017) gebruikt. Omdat alle schaalcores bij benadering al normaal verdeeld waren, volstond een lineaire transformatie naar T-scores (gegevens op te vragen bij de eerste auteur).

Behandeluitkomst

Allereerst maten we de behandeluitkomst af aan de verschilscore tussen de voormeting en de hermeting. Dit is een goede indicator van wat in het meetinterval bereikt is, zij het dat zo'n verschilscore een nogal abstracte maat is (De Beurs e.a. 2016). Een klinisch meer aansprekende indicatie van de behandeluitkomst is of de patiënt betrouwbaar is veranderd (meer veranderd dan toe te schrijven aan meeton nauwkeurigheid van het instrument) of zelfs is hersteld. Jacobson e.a. (1986, 1999) hebben criteria opgesteld voor betrouwbare verandering en herstel, die breed ingang hebben gevonden (Wise 2004; Sloan e.a. 2005). Toegepast op T-scores kozen wij voor een verschilscore > 5 voor betrouwbare verandering en T < 5 voor herstel (De Beurs e.a. 2019a; b).

Statistische analyse

De correlatie tussen de PCL-5- en de OQ-45-scores bij voormeting, bij hermeting en voor de pre-hermetingsverschilsscores werd berekend. De gemiddelde voor- en hermetingscore werd bepaald en per meetinstrument (en subschaal) werd met een t-toets voor gepaarde waarnemingen het verschil over de tijd getoetst. De effectgrootte werd uitgedrukt in Cohens d ($M_{pre} - M_{post} / \text{pooled SD}$). Dit gaf een eerste indruk van onderlinge verschillen in responsiviteit van (sub)schalen.

Voormeting- en hermetingscores op de PCL-5 en de OQ-45 werden ook vergeleken met variantieanalyse voor herhaalde metingen, in navolging van eerdere vergelijkingen van responsiviteit bij generieke meetinstrumenten (De Beurs e.a. 2012), bij depressie (De Beurs e.a. 2019c), angststoornissen (Schawo e.a. 2019) en eetstoornissen (Dingemans & Van Furth 2017). Bij deze analyse werd getoetst of er: 1. verschil was in score op de twee meetinstrumenten, 2. verschil in score over de tijd en 3. de interactie tussen meetinstrument en tijd. Een significante interactie betekent dat de hellingshoek van verandering over de tijd verschilt tussen beide instrumenten.

Ten slotte onderzochten we met een χ^2 -toets of de twee instrumenten overeenstemden in proporties patiënten die betrouwbaar waren verbeterd of waren hersteld. De mate van overeenstemming werd uitgedrukt in Spearmans rangdecorrelatiecoëfficiënt.

RESULTATEN

Eerst bepaalden we om de convergente validiteit van de PCL-5 en de OQ-45 na te gaan de correlatie tussen beide schalen, zowel op voor- als hermetingscores (TABEL 1) als op de verschilscore tussen voor- en hermeting (TABEL 2). Alle correlatiecoëfficiënten waren statistisch significant. De correlatie tussen schalen die min of meer hetzelfde meten (bijvoorbeeld de totaalscore van de PCL-5 en de OQ-SD-

TABEL 1 Pearsons product-momentcorrelatie (ppmc) tussen ruwe scores op de OQ-45 en de PCL-5; rechterbovenhoek is voormeting, linkeronderhoek is nameting

Correlatie	1	2	3	4	5	6	7	8
1. PCL-5 TOT ¹		0,86	0,66	0,89	0,87	0,76	0,77	0,77
2. Cluster B	0,91		0,54	0,63	0,67	0,63	0,63	0,75
3. Cluster C	0,81	0,72		0,51	0,45	0,42	0,42	0,49
4. Cluster D	0,94	0,77	0,70		0,69	0,73	0,72	0,61
5. Cluster E	0,93	0,79	0,71	0,82		0,65	0,70	0,70
6. OQ-TOT ²	0,87	0,74	0,65	0,86	0,82		0,95	0,85
7. OQ-SD ¹	0,88	0,77	0,65	0,86	0,84	0,97		0,92
8. OQ-ASD ²	0,87	0,80	0,70	0,79	0,83	0,89	0,94	

¹ n = 464

² n = 163

schaal), was aanzienlijk. De correlatie was bij de voormeting $r = 0,77$, bij de hermeting $r = 0,88$ en de correlatie van de verschillen was $r = 0,75$. Alleen de scores op cluster C van de PCL-5 ('vermijden') correleerden bij de voormeting wat lager met de andere scores ($r = 0,42$ tot $r = 0,51$). **TABEL 2** laat hetzelfde beeld zien met iets minder samenhang van verandering op de cluster C-schaal met de rest.

In **TABEL 3** is de gemiddelde T-score (met SD) weergegeven van de patiënten van wie twee scores per instrument beschikbaar waren, samen met de uitkomst van t-tests. De PCL-5-totaalscore bleek het gevoeligst voor verandering ($d = 0,49$), maar de OQ-SD deed daar in responsiviteit niet veel voor onder ($d = 0,44$). Het scoreverloop op beide instrumenten is weergegeven in **FIGUUR 1**. De gemiddelde T-score op de twee instrumenten bij voormeting verschilde aanzienlijk, maar het verschil in responsiviteit was gering (Delta T = 5,0 vs. 4,5 voor respectievelijk de PCL-5 en de OQ-SD).

In **FIGUUR 2** is de samenhang tussen voor- en hermetingverschilscores (Delta T) op de PCL-5-totaalscore en op de OQ-SD-schaal weergegeven; voor de meeste respondenten (twee derde) was het verschil in scoreafname op beide instrumenten < 6 T-punten (deze bevinden zich tussen de twee diagonale lijnen).

Vervolgens vergeleken we de uitkomsten op de PCL-5 en de OQ-45 bij voor- en hermeting met variantieanalyse voor herhaalde metingen. Voor 163 respondenten met een herhaalde meting op de PCL-5, de OQ-TOT en de OQ-SD werd een ANOVA voor herhaalde metingen uitgevoerd in een 2-bij-3-opzet (2 tijdstippen en 3 instrumenten). De resultaten van de toetsing gaven een significant tijdseffect: $F(1,162) = 30,69$; $p < 0,001$; $\eta^2 = 0,159$, een significant instrumenteffect voor het contrast tussen de PCL-5 en de twee OQ-schalen: OQ-TOT: $F(1,162) = 162,13$; $p < 0,001$; $\eta^2 = 0,500$ en OQ-SD: $F(1,162) = 150,23$; $p < 0,001$; $\eta^2 = 0,481$ (de OQ-schalen onder-

TABEL 2 Correlatie (ppmc) tussen pre-postverschilscores op de OQ-45 en de PCL-5

Correlatie	2	3	4	5	6	7	8
1. PCL-5 TOT ¹	0,85	0,63	0,92	0,86	0,77	0,75	0,75
2. Cluster B		0,46	0,68	0,64	0,65	0,61	0,68
3. Cluster C			0,49	0,45	0,34	0,42	0,35
4. Cluster D				0,73	0,73	0,71	0,68
5. Cluster E					0,63	0,65	0,62
6. OQ-TOT ²						0,96	0,86
7. OQ-SD ¹							0,92
8. OQ-ASD ²							

¹ n = 464

² n = 166

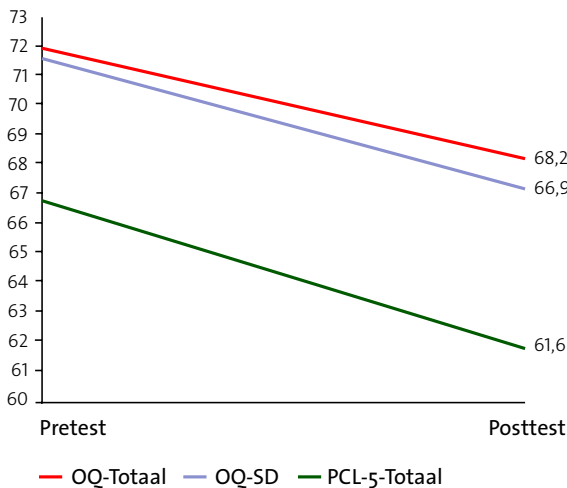
TABEL 3 Gemiddelde T-score (en SD) van patiënten met een voor- en nameting op de PCL-5 en patiënten met een voor- en nameting op de OQ-45, correlatie, t-waarden en binnen-subjectseffectgrootte (Cohens d)

T-scores	Voormeting		Nameting		r	t ¹	Cohens d ²
	M	SD	M	SD			
PCL-5-totaalscore	66,61	9,08	61,59	11,41	0,61	11,68	0,49
Cluster B	62,75	10,26	58,13	11,36	0,57	9,90	0,43
Cluster C	58,61	9,27	55,37	9,91	0,47	7,07	0,35
Cluster D	65,58	8,96	60,72	10,90	0,55	10,93	0,49
Cluster E	66,72	8,98	62,66	10,87	0,63	10,01	0,42
OQ-TOT	71,85	10,17	68,19	12,43	0,63	4,72	0,32
OQ-SD	71,44	9,04	66,93	11,42	0,65	10,99	0,44
OQ-IR	66,89	11,31	64,37	13,00	0,59	2,87	0,21
OQ-SR	68,01	11,55	64,13	13,63	0,55	4,07	0,31
OQ-ASD	71,78	12,01	68,37	13,82	0,70	4,27	0,26

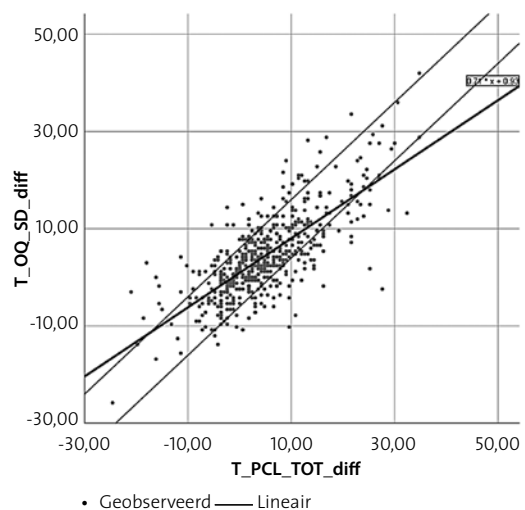
¹ Uitkomst van gepaarde t-toets; alle $p < 0,001$

² Cohens $d = \frac{M_{pre} - M_{post}}{\text{pooled SD}}$

FIGUUR 1 Scoreverloop op de OQ-45-totaalscore, de OQ-SD-schaal en de PCL-5



FIGUUR 2 Scatterplot voor de relatie tussen verschillcores op de PCL-5-totaalscore en de OQ-SD-schaal (beide na lineaire transformatie in T-scores)



ling verschillen niet). Echter, de interactie van tijd en instrument was niet significant voor het contrast tussen PCL en OQ-TOT: $F(1,162) = 1,60$; $p = 0,207$; $\eta^2 = 0,010$ en niet voor het contrast tussen PCL en OQ-SD: $F(1,162) = 0,86$; $p = 0,355$; $\eta^2 = 0,005$.

Met **FIGUUR 1** illustreren we de gevonden effecten duidelijk. T-scores op de beide OQ-schalen waren ongeacht het tijdsmoment hoger, alle scores namen af en op alle drie de schalen in gelijke mate. Het aflopen van de lijnen voor de verschillende instrumenten illustreert het tijdseffect en dit verliep parallel (geen interactie tussen tijd en instrument). Voor een paarsgewijze vergelijking van alleen de OQ-SD-

schaal met de PCL-5-totaalscore waren veel meer data beschikbaar ($n = 464$) en die analyseerden we ook nog apart, hetgeen hetzelfde resultaat opleverde: een significant tijdseffect: $F(1,463) = 147,14$; $p < 0,001$; $\eta^2 = 0,241$, een significant instrumenteffect $F(1,463) = 502,82$; $p < 0,001$; $\eta^2 = 0,521$ en wederom was de interactie van tijd en instrument niet significant voor de vergelijking tussen PCL-5 en OQ-SD: $F(1,463) = 2,80$; $p = 0,095$; $\eta^2 = 0,006$. Kortom: beide instrumenten gaven een gelijke verandering aan en waren dus even responsief.

TABEL 4 Betrouwbare verandering (reliable change index; RCI) en klinisch relevante verandering (clinical significance; CS) van patiënten met zowel een PCL-5 en OQ-SD en in 4 uitkomstcategorieën

T-scores	RCI		CS		Verslechterd		Onveranderd		Verbeterd		Hersteld	
	N	%	N	%	n	%	n	%	n	%	n	%
PCL-5 Totaal	203	43,8	93	20,0	41	8,8	220	47,4	111	23,9	92	19,8
OQ-SD	187	40,3	61	13,1	54	11,6	233	48,1	132	28,4	55	11,9

TABEL 4 geeft voor de PCL-5- en de OQ-SD-schaal de behandeluitkomst in de klinische betekenisvolle categorisering volgens Jacobson e.a. (1986, 1999) weer. Onder RCI staat het aantal patiënten (en percentage van het totaal) met betrouwbare verandering; onder CS het aantal patiënten met klinisch betekenisvolle verandering. De proporties verschilden voor beide indicatoren ten gunste van de PCL-5: $\chi^2(df = 1; n = 464) > 66,70; p < 0,001$. Beide indicatoren gecombineerd leidden tot de vier uitkomstcategorieën: ‘verslechterd’ (verslechtering en terugval samengenomen), ‘onveranderd’, ‘verbeterd’ en ‘hersteld’. De categorisering volgens de PCL-5 en volgens de OQ-SD was significant verschillend: $\chi^2(df = 9; n = 464) = 241,12; p < 0,001$; de behandeluitkomst volgens de PCL-5 was ook hier gunstiger dan volgens de OQ-SD. Er was wel een significante overeenkomst in categorisering (Spearman's rangordecorrelatiecoëfficiënt) volgens de PCL en de OQ-SD van $\rho = 0,63 (p < 0,01)$.

DISCUSSIE

De belangrijkste bevindingen van dit onderzoek zijn dat de PCL-5-totaalscore en de OQ-SD-score hoog correleren en vergelijkbaar responsief zijn. De sterke samenhang in score op beide instrumenten is extra opvallend wanneer we bedenken dat de periode waarover de respondent bevroegd wordt, verschilt (bij de PCL de afgelopen maand, bij de OQ de afgelopen week) en dat de antwoordopties ook verschillen (de PCL vraagt naar de hoeveelheid last, de OQ vraagt naar frequentie van voorkomen). De PCL-5 en de OQ-SD geven nagenoeg hetzelfde resultaat aan: de effectgrootte van het test-hertestverschil is volgens beide instrumenten $d = 0,49$ en $0,44$ (en $\Delta T = 5,0$ en $4,5$), voor respectievelijk de PCL-5 en de OQ-SD.

De resultaten van de variantieanalyse voor herhaalde metingen bevestigen dit: zelfs met een grote statistische power vonden we geen verschil tussen de PCL-5 en de OQ-SD in responsiviteit: de hoogte van de scores verschilde, maar de hellingshoek van afname in score was gelijk. Wanneer de instrumenten worden ingezet om op groepsniveau uitspraken te doen over het behaalde therapie-effect, dan doen ze dus nauwelijks voor elkaar onder in responsiviteit. Vergelijken we de proporties verbeterde en herstelde

patiënten volgens de PCL-5 en de OQ-SD, dan vinden we wel enig verschil. De PCL-5 gaf een gunstiger beeld van de behandeluitkomst dan de OQ-SD met meer herstelde patiënten. Dit verschil in uitkomst kan veroorzaakt zijn door in het algemeen hogere T-scores op de OQ-SD dan op de PCL-5 (zie ook **FIGUUR 1**), waardoor op de OQ-SD een geringer aantal respondenten de grensscore van $T = 55$ kon overschrijden. Dit kan het gevolg zijn van gebruik van inadequate Canadese normgegevens voor de T-scoreconversie van de PCL-5. Verhoging van de T-scores op de PCL-5 met 5 punten geeft resultaten (13,6% hersteld) die veel beter overeenstemmen met de OQ (11,9% hersteld); we pleiten dan ook voor het verzamelen van goede normgegevens voor de PCL-5 bij een Nederlandse bevolkingssteekproef.

Maar uiteenlopende resultaten op de PCL-5 en OQ-SD kunnen ook een correcte afspiegeling zijn van verschillen in meetpretentie tussen de instrumenten. Iemand kan herstellen volgens de OQ, maar toch nog zoveel last hebben van herbelevingen volgens de PCL-5 dat van herstel nog geen sprake is. Of omgekeerd: iemand is hersteld volgens de PCL-5, maar er zijn nog te veel generieke klachten volgens de OQ om te spreken van herstel.

Sterke punten en beperkingen

Een sterk aspect van dit onderzoek is dat we een flink databestand hebben geanalyseerd en daarmee over voldoende statistische power beschikten om zelfs een klein verschil in responsiviteit aan te tonen. De gegevens zijn verzameld in de alledaagse praktijk van het Sinai Centrum, waardoor de bevindingen goed te generaliseren zijn naar de klinische werkelijkheid van een kliniek gespecialiseerd in de behandeling van PTSS.

Een beperking aan het onderzoek is dat de hermeting bij veel patiënten plaatsvond tijdens hun behandeling en soms al vroeg in de behandeling, bijvoorbeeld al na vier weken, waardoor niet het effect van het gehele zorgtraject is vastgesteld. Een aanvullende analyse van alleen de patiënten met meer dan 20 weken tussen de twee metingen ($n = 346$; 74,6% van de gehele groep met een hermeting) wees uit dat de ΔT op de PCL-5 dan toenam van 5,0 naar 6,1 (SD: 9,6).

Gemiddeld minder verandering meten verkleint de kans om een verschil in responsiviteit tussen de PCL-5 en de OQ-SD aan te tonen, al wordt dit wel gecompenseerd door de totale omvang van de onderzochte groep. Bij volledige zorgtrajecten mogen we een Delta T van rond de 8,0 verwachten en dan zal het verschil tussen de PCL en de OQ-SD ongeveer één Delta T-punt zijn. Hiermee zou de PCL-5 bijna 10% meer responsief zijn dan de OQ-SD.

De toepassing van de PCL-5 in de klinische praktijk bleek nuttig, want de PCL-5 biedt informatie over de ernst van de PTSS en de score op subschalen levert per patiënt een profiel op voor de aard van de aandoening op basis van de symptoomclusters: staan bij deze patiënt de herbeleving, de vermijding, de negatieve gedachten en stemming of de arousal en snelle schrikreacties op de voorgrond? Deze informatie is bruikbaar om de behandeling richting te geven, te ondersteunen en te evalueren.

De voordelen van gebruik van de PCL-5 moeten wel opwegen tegen de extra inspanning die gevraagd wordt van de patiënt (beantwoording van 20 extra vragen). Een ander nadeel van gebruik van stoornisspecifieke instrumenten is de beperkte aggregaerbaarheid van de meetresultaten voor kwaliteitsmanagement. De behandeluitkomst volgens de PCL-5 is niet bruikbaar op een hoger aggregatieniveau dan de groep met PTSS: de uitkomst is immers niet zonder meer vergelijkbaar met de uitkomst van een ander instrument bij patiënten met andere stoornissen, zoals

depressie of paniekstoornis. Voor een instituut als Sinai, gespecialiseerd in de behandeling van PTSS, kan het een passend instrument zijn, maar een vrijgevestigde psychiater met diverse patiënten in de praktijk krijgt met een waaier van stoornisspecifieke meetinstrumenten geen goed zicht op eigen behandelresultaten.

Ten slotte: een psychiatrische aandoening is meer dan een set symptomen en volgens de herstelvisie moet behandeling niet alleen gericht zijn op symptoomreductie, maar ook op maatschappelijk en persoonlijk herstel (Brouwers e.a. 2013) en moet er voor een complete uitkomstmeting ook het functioneren en kwaliteit van leven gemeten worden. Er moet nader worden onderzocht of zo'n verbreding van de behandeluitkomst tot betere ROM-informatie leidt en voldoende opweegt tegen de extra inspanning die het van de patiënt vergt.

CONCLUSIE

Qua responsiviteit ontlopen de OQ-SD en de PCL-5 elkaar niet en voor het evalueren van behandeluitkomsten op geaggregeerd niveau (kwaliteitsmanagement) zijn beide instrumenten in principe bruikbaar. Aangezien de PCL-5 beter bij de doelgroep past, meer gedetailleerde informatie biedt om de behandeluitkomst te evalueren en enigszins gevoeliger is om herstel of betrouwbare verandering te detecteren, raden we aan bij PTSS de ROM-set aan te vullen met de PCL-5.

LITERATUUR

- Aarts I, Snoek A, Vriend C, van den Heuvel O, Thomas K. PROSPER: Prediction and Outcome Study of PTSD and Personality Disorders: Design of Prediction and Mediation Study. *Eur J Psychotraumatol* 2019; 10.
- American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-5)*. Washington: APA; 2013.
- Ashbaugh AR, Houle-Johnson S, Herbert C, El-Hage W, Brunet A. Psychometric validation of the English and French versions of the posttraumatic stress disorder checklist for DSM-5 (PCL-5). *PLoS One* 2016; 11(10): e0161645.
- Blevins CA, Weathers FW, Davis MT, Witte TK, Domino JL. The posttraumatic stress disorder checklist for DSM-5 (PCL-5): Development and initial psychometric evaluation. *J Traum Stress* 2015; 28: 489-98.
- Boeschoten M, Bakker A, Jongedijk R, Olff M. PTSS checklist voor de DSM-5 (PCL-5). 2014. Unpublished manuscript.
- Brouwers E, van Gestel-Timmermans H, van Nieuwenhuizen C. Herstelgerichte zorg in Nederland: modegril of daadwerkelijke verandering? *Psychopraktijk* 2013; (5): 14-7.
- de Beurs E, Barendregt M, de Heer A, van Duijn E, Goeree B, Kloos M, e.a. Comparing methods to denote treatment outcome in clinical research and benchmarking mental health care. *Clin Psychol Psychother* 2016; 23: 308-18.
- de Beurs E, Barendregt M, Flens G, van Dijk E, Huijbrechts I, Meerding JW. Vooruitgang in de behandeling meten: Een vergelijking van vragenlijsten voor zelfrapportage. *Maandblad Geestelijke Volksgezondheid* 2012; 67: 259-70.
- de Beurs E, Barendregt M, Warmerdam L, red. *Behandeluitkomsten: bron voor kwaliteitsbeleid in de GGZ*. Amsterdam: Boom; 2017.
- de Beurs E, Carlier IV, van Hemert AM. Approaches to denote treatment outcome: Clinical Significance and Clinical Global Impression compared. *Int J Meth Psychiatr Res* 2019b; 28.
- de Beurs E, den Hollander-Gijsman M, Buwalda V, Trijsburg W, Zitman FG. De Outcome Questionnaire (OQ-45): een meetinstrument voor meer dan alleen psychische klachten. *De Psycholoog* 2005; 40(1): 53-63.

- de Beurs E, den Hollander-Gijsman ME, van Rood YR, van der Wee NJ, Giltay EJ, van Noorden MS, e.a. Routine outcome monitoring in the Netherlands: Practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clin Psychol Psychother* 2011; 18: 1-12.
- de Beurs E, Flens G. Gebruik van verschillende meetinstrumenten: de getransformeerde T-score en equivalente responsiviteit. In: de Beurs E, Warmerdam E, Barendregt M, red. *Behandeluitkomsten: bron voor kwaliteitsbeleid in de GGZ*. Amsterdam: Boom; 2017. p. 225-38.
- de Beurs E, Flens G, Williams G. Meetresultaten interpreteren in de klinische psychologie: een aantal voorstellen. *De Psycholoog* 2019a; 54(6): 10-23.
- de Beurs E, Vissers E, Schoevers R, Carlier IVE, van Hemert AM, Meesters Y. Comparative responsiveness of generic versus disorder-specific instruments for depression: An assessment in three longitudinal datasets. *Depress Anxiety* 2019c; 36: 93-102.
- de Beurs E, Zitman FG. De Brief Symptom Inventory (BSI): De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90. *Maandblad Geestelijke Volksgezondheid* 2006; 61: 120-41.
- de Jong K, Nugter MA, Polak MG, Wagenborg JEA, Spinhoven P, Heiser WJ. (2007). The Outcome Questionnaire (OQ-45) in a Dutch population: A cross-cultural validation. *Clin Psychol Psychother* 2007; 14: 288-301.
- Dingemans AE, van Furth EF. Het meten van verandering tijdens behandeling voor eetstoornissen: Een vergelijking van twee algemene en specifieke vragenlijsten. *Tijdschr Psychiatr* 2017; 59: 278-85.
- Fairburn CG, Beglin SJ. *Eating Disorder Examination Questionnaire (EDE-Q 6.0)*. In: fairburn CG, red. *Cognitive Therapy and Eating Disorders*. New York: Guilford Press; 2008. p. 309-13.
- Hermann RC. *Improving mental healthcare: a guide to measurement-based quality improvement*. Washington: American Psychiatric Publishing; 2007.
- Jacobson NS, Follette WC, Revenstorf D. (1986). Toward a standard definition of clinically significant change. *Behav Ther* 1986; 17: 308-11.
- Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB. Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. *J Consult Clin Psychol* 1999; 67: 300-7.
- Kilbourne AM, Beck K, Spaeth-Rublee B, Ramanuj P, O'Brien RW, Tomoyasu N, e.a. Measuring and improving the quality of mental health care: a global perspective. *World Psychiatry* 2018; 17: 30-8.
- Lambert MJ, Burlingame GM, Umphress VJ, Hanssen N, Vermeersch D, Clause G, e.a. The reliability and validity of the Outcome Questionnaire. *Clin Psychol Psychother* 1996; 3: 249-58.
- McCall WA. *How to measure in education*. New York, MacMillan; 1922.
- Nugter MA, Hermens ML, Robbers S, Van Son G, Theunissen J, Engelsbel F. Use of outcome measurements in clinical practice: How specific should one be? *Psychother Res* 2019; 29: 432-44.
- Schawo SJ, Carlier IVE, van Hemert AM, de Beurs. Measuring treatment outcome in patients with anxiety disorders: A comparison of the responsiveness of generic and disorder-specific instruments. *J Anxiety Disord* 2019; 64: 55-63.
- Sloan JA, Cella D, R. D. Hays RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. *J Clin Epidemiol* 2005; 58: 1217-9.
- Timman R, de Jong K, de Neve-Enthoven N. Cut-off scores and clinical change indices for the Dutch Outcome Questionnaire (OQ-45) in a large sample of normal and several psychotherapeutic populations. *Clin Psychol Psychother* 2017; 24: 72-81.
- van der Mheen M, ter Mors LM, van den Hout MA, Cath DC. Routine outcome monitoring bij de behandeling van angststoornissen: diagnosespecifieke versus generieke meetinstrumenten. *Tijdschr Psychiatr* 2018; 60: 11-9.
- Weathers FW, Litz BT, Herman DS, Huska JA, Keane TM. *The PTSD Checklist (PCL): Reliability, validity, and diagnostic utility*. Annual Convention of the International Society for Traumatic Stress Studies. San Antonio; 1993.
- Weathers FW, Litz BT, Keane TM, Palmieri PA, Marx BP, Schnurr PP. *The PTSD Checklist for DSM-5 (PCL-5)*. 2013. <https://www.ptsd.va.gov/professional/assessment/adult-sr/ptsd-checklist.asp>.
- Wise EA. Methods for analyzing psychotherapy outcomes: A review of clinical significance, reliable change, and recommendations for future directions. *J Pers Assess* 2004; 82: 50-9.

SUMMARY

The PTSD Checklist for DSM-5 (PCL-5): comparing responsivity with the Outcome Questionnaire (OQ-45) and practical utility

E. DE BEURS, K. THOMAES, H. KRONEMEIJER, J. DEKKER

BACKGROUND The PTSD Checklist for the DSM-5 (PCL-5) may be a suitable addition for routine outcome monitoring (ROM) for patients with PTSD.

AIM To determine whether the PCL-5 is worth the extra effort that administration requires from the patient.

METHOD Pretest and retest measurement results of the PCL-5 and the OQ-45 were compared head-to-head in 464 patients from the Sinai Center of Arkin.

RESULTS The correlations between scores on the instruments were high and analysis of variance for repeated measurements revealed no difference in responsiveness. Comparison of Cohen's d (0.49 vs. 0.43) and Delta T (5.0 vs. 4.4), indicated a slightly better responsiveness of the PCL-5 and also the proportion of recovered patients was greater according to the PCL-5 compared to the OQ-SD.

CONCLUSION At first glance, the PCL-5 and the OQ-SD were equally sensitive to detect change during treatment. However, the PCL-5 provided more detailed information about the nature and severity of symptomatology in an individual patient and with the PCL-5 we were slightly better able to demonstrate clinical significant change than with the OQ-SD. We recommend to add the PCL-5 to ROM for patients with PTSD.

TIJDSCHRIFT VOOR PSYCHIATRIE 62(2020)6, 448-456

KEYWORDS OQ-45, PCL-5, post-traumatic stress disorder, responsivity, ROM