

Inventory of Depressive Symptomatology en verkorte versie in routine outcome monitoring van Stichting Benchmark GGZ

Y. MEESTERS, W.B. DUIJZER, W.A. NOLEN, R.A. SCHOEVERS, H.G. RUHÉ

ACHTERGROND Verzekeraars vragen in toenemende mate een getalsmatige onderbouwing van de behandelresultaten in de ggz. Daarbij baseren zij zich op de door de Stichting Benchmark GGZ (SBG) aangeleverde data en procedures. SBG heeft een beperkt aantal (generieke) vragenlijsten geaccepteerd om te kunnen benchmarken. Binnen de klinische praktijk voldoen die instrumenten niet om daarmee routine outcome monitoring (ROM) mee te kunnen verrichten. De behoefte bestaat om daar een veelgebruikte depressievragenlijst, de Inventory of Depressive Symptomatology (IDS), voor te gebruiken.

DOEL Argumenten aandragen waarom de IDS een uitstekend en bruikbaar instrument is om door SBG als meetinstrument te worden geaccepteerd.

METHODE Door middel van een literatuuroverzicht de kwaliteit van de IDS bespreken.

RESULTATEN De IDS is een kwalitatief goed instrument om de ernst van depressies te meten en is tevens gevoelig om verbetering in de stemming te meten.

CONCLUSIE De IDS is een uitstekend instrument om ROM-metingen mee te verrichten en is zonder kosten verkrijgbaar. Er wordt voor gepleit dat SBG deze vragenlijst accepteert als een van de instrumenten om te kunnen benchmarken.

TIJDSCHRIFT VOOR PSYCHIATRIE 58(2016)1, 48-54

TREFWOORDEN benchmarking, IDS, ROM, SBG



ARTIKEL



Er is veel discussie over de toepassing van *routine outcome monitoring* (ROM) en *benchmarking* zoals die door de Stichting Benchmark GGZ (SBG) en zorgverzekeraars wordt voorgestaan. De kritiek betreft de wetenschappelijke onderbouwing, medisch-ethische aspecten, maar ook de uitvoerbaarheid in de praktijk en de belasting voor patiënten (Hafkenscheid & Van Os 2013; Van Os e.a. 2012).

Waar de partijen het over eens zijn, is dat ROM zinvol is om de voortgang van behandelingen in de praktijk te monitoren, deze samen met de patiënt te evalueren, en de behandeling zo nodig bij te stellen. Het is dan van belang om instrumenten te kiezen die in die praktijk ook zinvol en

acceptabel zijn voor patiënten en behandelaars, en die daadwerkelijk een score opleveren die relevant is om het beloop van een specifieke behandeling te evalueren. Bovendien is het voor een zinvolle toepassing uiterst plezierig als het instrument zich al in de praktijk heeft bewezen (Hafkenscheid & Van Os 2014).

Dit alles geldt niet voor de *Outcome Questionnaire 45* (OQ45; De Jong e.a. 2009), een instrument dat op dit moment door zeer veel patiënten wordt ingevuld, maar waarvan de validiteit nog onvoldoende is aangetoond en goede normering ontbreekt (Egberink e.a. 2009-2014). Daarbij is van het flinke aantal items bijna een kwart niet van toepassing

voor psychiatrische patiënten die bijvoorbeeld geen werk (meer) hebben, geen opleiding hebben kunnen afronden en geen seksuele relatie hebben (Hafkenscheid & Van Os 2014). Benchmarking in de ggz is op zich al niet valide en bij de keuze van de meetinstrumenten door SBG zetten wij eveneens vraagtekens. Het lijkt alsof bij de keuze van de vragenlijsten het benchmarken belangrijker wordt gevonden dan de vraag of meetinstrumenten in de dagelijkse praktijk bruikbaar zijn (Van Os e.a. 2012).

Werkwijze SBG

SBG heeft de ggz in zorgdomeinen onderverdeeld. Eén daarvan is ‘volwassenen kort’ waarbij de behandeling vooral gericht is op herstel of op het verminderen van klachten en symptomen. Een groot deel van de patiënten die onder dit zorgdomein vallen, heeft een depressieve stoornis. Volgens het NEMESIS-2-onderzoek blijkt van alle psychische aandoeningen in Nederland de incidentie van depressie het hoogst te zijn (De Graaf e.a. 2012). Opmerkelijk is dat het door SBG goedgekeurde instrumentarium geen specifieke, in de reguliere ggz gebruikte, vragenlijst voor het meten van depressies bevat, maar vooral zogenaamde generieke instrumenten. Dat zijn instrumenten die geestelijke gezondheidsklachten zouden meten zonder daarbij te differentiëren naar diagnose of klachten.

SBG is er voorstander van dat men uiteindelijk één generiek instrument zal kiezen waarmee men alle behandelingen in de ggz kan vergelijken om te kunnen benchmarken. Dat lijkt ons een onterechte simplificatie en reductie van psychiatrische problematiek. Wil men behandelingen met elkaar kunnen vergelijken, dan zal men op zijn minst moeten weten waarvoor behandeld wordt, temeer omdat juist deze uitkomsten voor patiënt en behandelaar ook in de spreekkamer relevant zijn. Om de vergelijking met de somatiek te maken: een behandeling van hypertensie verschilt wezenlijk van die van een maligniteit. Hetzelfde geldt voor autisme, eetstoornissen en depressie.

Specifieke instrumenten

Wij pleiten dan ook voor het gebruik van meer specifieke klachtgerichte instrumenten, die gecombineerd zullen moeten worden met diagnosegegevens en maten voor chroniciteit en/of recidivering om zinvolle informatie op te leveren. Omdat de wetenschappelijke raad van SBG aanbeveelt het instrumentarium werkende weg te willen verbeteren (Blijd-Hoogewys e.a. 2013), doen wij bij dezen de aanbeveling de IDS als specifieke maat voor depressiviteit toe te voegen aan het instrumentarium om zo tot een wezenlijke verbetering te komen.

De *Inventory of Depressive Symptomatology* (IDS) en de afgeleide, verkorte versie, de Quick-IDS (QIDS), voldoen beide aan een aantal voor praktijk en wetenschap belangrijke crite-

AUTEURS

YBE MEESTERS, klinisch psycholoog/psychotherapeut, Rijksuniversiteit Groningen, UMC Groningen, afd. Psychiatrie, Stemmings- en Angststoornissen, Groningen.

WIANNE B. DUIJZER, GZ-psycholoog, Rijksuniversiteit Groningen, UMC Groningen, afd. Psychiatrie, afd. Algemene Psychiatrie, Groningen.

WILLEM A. NOLEN, psychiater, emeritus hoogleraar Psychiatrie, Groningen.

ROBERT A. SCHOEVERS, psychiater, hoogleraar Psychiatrie en hoofd van de afd. Psychiatrie, UMC Groningen.

HENRICUS G. RUHÉ, psychiater, Rijksuniversiteit Groningen, UMC Groningen, afd. Psychiatrie, Stemmings- en Angststoornissen, Groningen.

CORRESPONDENTIEADRES

Dr. Y. Meesters, Universitair Centrum Psychiatrie, UMC Groningen, Postbus 30.001, 9700 RB Groningen.
E-mail: y.meesters@umcg.nl

Strijdige belangen: W. Nolen was betrokken bij de vertaling van de IDS en zijdelings bij de validatie. H.G. Ruhé meldde grants van ZonMw VENI (#016.126.059), een Investigator Initiated Trial Grant van Lundbeck, en sprekershonoraria van Lundbeck en Astra Zeneca, relevant buiten dit artikel. Y. Meesters is lid van de Expertraad Volwassenenen Kort van Stichting Benchmark GGZ.

Het artikel werd voor publicatie geaccepteerd op 8-4-2015.

ria. Het is namelijk een valide en betrouwbaar instrument dat gevoelig is voor het meten van veranderingen en eenvoudig af te nemen. Het is niet goed uit te leggen dat gebruik van dit instrument in de praktijk niet door de zorgverzekeraar wordt erkend. In dit artikel zullen wij toelichten waarom in onze ogen de IDS een uitstekende kandidaat is om op zeer korte termijn te worden opgenomen in het door SBG gehanteerde ROM-instrumentarium.

Het meetinstrument

De IDS werd ontwikkeld om depressie bij opgenomen en ambulante patiënten te meten. In de vragenlijst zijn alle criteria van de DSM-classificaties voor depressieve stemmingsstoornissen opgenomen en geoperationaliseerd (Rush e.a. 1986, 1996). De behoefte aan een nieuwe vragenlijst ontstond mede omdat andere, internationaal veel gebruikte vragenlijsten om de ernst van depressie te meten, zoals de *Hamilton Depression Rating Scale* (HDRS), de *Montgomery Åsberg Depression Rating Scale* (MADRS) en de *Beck Depression Inventory* (BDI), onvoldoende ‘dekking’ gaven van alle DSM-items en subclassificaties.

FIGUUR 1 Zelfinvullijst depressieve symptomen

ZELFINVULLIJST DEPRESSIEVE SYMPTOMEN
(INVENTORY OF DEPRESSIVE SYMPTOMATOLOGY: IDS-SR)¹
(In te vullen door patiënt)

Naam:

Datum: - -

Kruis bij elke vraag het antwoord aan dat de afgelopen zeven dagen het meest op u van toepassing was

1. In slaap vallen:

- 0. Het duurt nooit langer dan 30 minuten om in slaap te vallen.
- 1. Het duurt tenminste 30 minuten om in slaap te vallen, minder dan de helft van de week.
- 2. Het duurt tenminste 30 minuten om in slaap te vallen, meer dan de helft van de week.
- 3. Het duurt meer dan 60 minuten om in slaap te vallen, meer dan de helft van de week.

2. Slaap gedurende de nacht:

- 0. Ik word 's nachts niet wakker.
- 1. Ik slaap onrustig en licht en word een aantal keren per nacht even wakker.
- 2. Ik ben tenminste één keer per nacht klaar wakker, maar val weer gemakkelijk in slaap.
- 3. Ik word vaker dan één keer per nacht wakker en blijf dan 20 minuten of langer wakker, meer dan de helft van de week.

3. Te vroeg wakker worden:

- 0. Meestal word ik niet eerder dan 30 minuten voordat ik op moet staan, wakker.
- 1. Ik word meer dan 30 minuten voordat ik op moet staan wakker, meer dan de helft van de tijd.
- 2. Ik word tenminste 1 uur voordat ik op moet staan wakker, meer dan de helft van de tijd.
- 3. Ik word tenminste 2 uur voordat ik op moet staan wakker, meer dan de helft van de tijd.

4. Te veel slapen:

- 0. Ik slaap niet langer dan 7-8 uur per nacht, zonder overdag een dutje te doen.
- 1. Ik slaap niet langer dan 10 uur binnen één etmaal (inclusief dutten).
- 2. Ik slaap niet langer dan 12 uur binnen één etmaal (inclusief dutten).
- 3. Ik slaap langer dan 12 uur binnen één etmaal (inclusief dutten).

5. Somber voelen:

- 0. Ik ben niet somber.
- 1. Ik ben minder dan de helft van de tijd somber.
- 2. Ik ben meer dan de helft van de tijd somber.
- 3. Ik ben bijna altijd somber.

¹ Nederlandse vertaling: Altrecht GGZ. Copyright © 1995/2005

In tegenstelling tot de andere meetinstrumenten worden in de IDS alle negen DSM-symptoomdomeinen van een depressie (depressieve stoornis) gemeten, evenals melancholische en atypische symptomen, alsmede vaak met depressie geassocieerde symptomen zoals angst en pijn (Rush e.a. 1986).

De IDS kent twee versies: een door een clinicus afgenomen gestandaardiseerd en gestructureerd interview (IDS-C) en een zelfbeoordelingsvragenlijst (IDS-SR). De structuur van de IDS-C en de IDS-SR is hetzelfde. De respondenten wordt standaard gevraagd hoe men zich de afgelopen 7 dagen heeft gevoeld. De IDS-C/SR bestaat uit 30 antwoorditems die

gescoord worden op een schaal van 0-3 (voor enkele voorbeeldvragen: zie **FIGUUR 1**). In de antwoorditems worden de ernst en frequentie meegenomen met duidelijke ijkpunten per item. De schaal van de IDS-C/SR varieert tussen 0-84. Iemand zonder klachten scoort niet hoger dan 9 op de IDS-C/SR. In internationaal onderzoek wordt een score < 12 (IDS-C) en < 14 (IDS-SR) aangehouden als een maat voor remissie van depressie.

Uit de 30-item-IDS-C/SR zijn verkorte versies ontwikkeld. De 16-item-QIDS (Rush e.a. 2003, 2005) kent ook een door de clinicus te scoren versie (QIDS-C) en een zelfbeoordelingsvorm voor patiënten (QIDS-SR). De uitkomsten van beide QIDS-versies correleren zeer hoog met de oorspronkelijke IDS-versies (Rush e.a. 1996). De schaal van de QIDS-C/SR loopt van 0-27. In internationaal onderzoek wordt een score < 6 op de QIDS-C/SR aangehouden als een remissie van depressie. De QIDS leent zich door de korte afnameduur en de overlap met de oorspronkelijke IDS uitstekend voor toepassing in ROM (Dingemans 2010).

De IDS-C/SR en QIDS-C/SR zijn zonder kosten beschikbaar en de Nederlandse versie is te downloaden via www.tijdschriftvoorpsychiatrie.nl/measuringinstruments. Daarnaast zijn veel relevante IDS-publicaties en de officiële vertaalde versies van de IDS beschikbaar via www.ids-qids.org.

Wetenschappelijke aspecten

De IDS is rond de tijd van de introductie uitgebreid geëvalueerd ten aanzien van validiteit (Bernstein e.a. 2006; Rush e.a. 1986, 1996, 2003; Trivedi e.a. 2004). Ook daarna heeft vervolgonderzoek de waarde van de IDS-C/SR en QIDS-C/SR verder onderbouwd (Schulte-Van Maaren e.a. 2013). Ze wordt in veel studies naar depressie gebruikt (zie o.a. Bernstein e.a. 2006, 2007; Brunoni e.a. 2013; Corruble e.a. 1999; Rush e.a. 1996; Trivedi e.a. 2004), inclusief de gezaghebbende Nederlandse Studie naar Depressie en Angst (NESDA, zie www.nesda.nl; o.a. Karsten e.a. 2010; Wardenaar e.a. 2012).

CONSTRUCTVALIDITEIT

Onder constructvaliditeit van een instrument verstaat men de mate waarin dat instrument in staat is daadwerkelijk te meten wat men beoogt te meten. Rush e.a. (1996) onderscheidde drie factoren in de IDS-C/SR: 1. cognitief/stemming, 2. angst/arousal, en 3. regulatie van slapen en eetlust. In een andere factoranalyse waarin IDS-, HDRS- en BDI-items werden gecombineerd, bleek dat de IDS de meeste symptomen van de gezamenlijke factorstructuur omvatte van de drie onderzochte instrumenten (Gullion e.a. 1998). De IDS is voldoende unifactorieel, wat inhoudt dat de algemene factor (ernst van depressie) sterk genoeg is zodat de aanwezigheid van eventuele kleinere factoren geen

invloed heeft op de meting van depressie (Wanders e.a. 2015). Daardoor zijn met een daartoe geëigende techniek (de item-responstheorie(IRT)-analyse, waarbij per vragenlijst de relatie tussen symptomen en de onderliggende ernst wordt gemodelleerd) conversietabellen ten opzichte van andere ernstvragenlijsten mogelijk (Rush e.a. 2003); zie ook www.ids-qids.org.

CONTENTVALIDITEIT

Bij de vraag naar contentvaliditeit wordt gekeken of een vragenlijst het gehele te meten begrip meet. De IDS-C/SR en de QIDS-C/SR omvatten alle negen DSM-symptoomdomeinen van de depressieve stoornis. Daarnaast bevat de IDS-C/SR aanvullende items voor melancholische en atypische kenmerken en items voor agitatie, angst en de zogenaamde endogene symptomen zoals eerder gedefinieerd in de *Research Diagnostic Criteria* (RDC; Spitzer e.a. 1978).

CONVERGERENDE VALIDITEIT

De convergerende validiteit beschrijft de correlatie van een vragenlijst met andere gevalideerde vragenlijsten.

De convergerende validiteit van de IDS-C met de HDRS wordt gerapporteerd tussen 0,68-0,95, met de BDI tussen 0,61-0,86, terwijl met de MADRS 0,81 is gerapporteerd. De IDS-SR correleert ook goed met de HDRS (0,67-0,88) en de BDI (0,76-0,93) (Corruble e.a. 1999; González e.a. 2014; Rush e.a. 1986, 1996). De correlatie tussen QIDS-SR en de HDRS was 0,81 (Rush e.a. 2004).

In diverse leeftijdsklassen werden eveneens goede resultaten gevonden, zoals bij jongeren (studenten van rond 20 jaar)(González e.a. 2013) en ouderen (Doraiswamy e.a. 2010). Ook in specifiekere populaties (bijvoorbeeld mensen met winterdepressie) is aangetoond dat de IDS-SR zowel hoog correleert met de BDI-II (0,93) als met een uitgebreide versie van de Hamilton-vragenlijst (SIGH-SAD-SR) (0,94) (Meininger 2007).

Wat betreft de correlatie tussen IDS- en QIDS-versies onderling: Rush e.a. (1996) vonden een correlatie tussen de IDS-C en de IDS-SR tussen 0,88-0,91 (Richtlijn Depressie 2013). De correlatie tussen QIDS-SR en IDS-SR ligt tussen 0,83-0,96 (Rush e.a. 2004; Trivedi e.a. 2004). Overigens waren de resultaten behaald met de QIDS-C in een studie bij ouderen, waarbij werd onderzocht of men met gebruik van de meetinstrumenten ouderen met en zonder depressie kon onderscheiden, iets beter dan de resultaten die werden vastgesteld met de QIDS-SR (Doraiswamy e.a. 2010). De beide varianten van de QIDS bleken ook bij patiënten met een bipolaire stoornis zeer goed bruikbaar (Bernstein e.a. 2010).

GEVOELIGHEID VOOR VERANDERING

De IDS-C/SR en de QIDS-C/SR zijn gevoelig gebleken voor verandering door behandeling met psychofarmaca, psycho-

therapie en andere behandelingen zoals elektroconvulsie-therapie. In vergelijkend onderzoek met de HDRS en de BDI is gebleken dat de IDS-C even gevoelig of zelfs gevoeliger is voor het meten van veranderingen in depressieve klachten (Biggs e.a. 2000; Corruble e.a. 1999; Rush e.a. 1996). Dit geldt ook voor de IDS-SR in vergelijking met de HDRS en de BDI (Rush e.a. 1986) en de MADRS en de SCL-90 (Corruble e.a. 1999). Ook ten opzichte van elkaar blijken de QIDS-SR en de IDS-SR even gevoelig om verandering te meten (Rush e.a. 2003; 2005). De onderlinge overeenstemming tussen IDS-C/SR en QIDS-C/SR om verandering te kunnen meten blijkt ook te bestaan bij patiënten met een bipolaire stoornis (Trivedi e.a. 2004).

INTERNE CONSISTENTIE

De mate waarin verschillende items van een vragenlijst waarmee men hetzelfde begrip beoogt te meten dat ook daadwerkelijk doen, wordt interne consistentie genoemd. Uit het oorspronkelijke onderzoek van Rush e.a. (1986) komt naar voren dat de IDS-SR een goede interne betrouwbaarheid heeft (Cronbachs $\alpha = 0,85$). Bij verder validatieonderzoek rapporteerden diverse onderzoekers (Biggs e.a. 2000; Corruble e.a. 1999; Rush e.a. 1996; 2003; Trivedi e.a. 2004) interne consistenties (Cronbachs α) voor de IDS-C tussen 0,67-0,94, voor de IDS-SR tussen 0,77-0,85, voor de QIDS-C 0,85 en voor de QIDS-SR 0,86. In een Nederlandse studie werd eveneens een goede interne consistentie (Cronbachs $\alpha = 0,78-0,94$) gevonden, behalve voor de subschaal die de atypische kenmerken meet (Cronbachs $\alpha = 0,68$) (Schulte-van Maaren e.a. 2013).

TRAINING

Voor de afname van de IDS-C en QIDS-C is minimale training noodzakelijk; de opgegeven ijkpunten zijn nuttig bij het scoren.

VALIDITEIT BUITEN HET DIAGNOSEGEBIED

Ook wanneer er op basis van de DSM geen depressieve stoornis kan worden gediagnosticeerd, kunnen mensen last hebben van klachten die passen bij somberheid. In het Nederlandse onderzoek van Karsten e.a. (2010) is de IDS-SR onderzocht betreffende de eigenschap een subklinische depressie te herkennen (een depressie die volgens het *Composite International Diagnostic Interview* (CIDI) niet voldoet aan de DSM-IV-criteria van de depressieve stoornis). Gebruikmakend van gegevens van 2157 deelnemers van NESDA bleek bij deze deelnemers een hogere score op de IDS-SR lineair geassocieerd te zijn met een toename van functionele beperkingen in het dagelijks leven. De IDS(-SR) is dus geen diagnostisch instrument, maar geeft wel een goede indruk van de ernst en de symptomen van een bre-

dere bandbreedte van stemmingsklachten van subklinische tot klinische depressie.

Wat betreft verstoring van de validiteit door somatische klachten bleek bij mensen met astma die tevens een depressie hadden dat de QIDS goed toepasbaar was om de ernst van depressieve klachten te meten (Sherwood e.a. 2008).

CONCLUSIE

Wij concluderen dat het voor de hand ligt om de vragenlijsten IDS-C/SR en QIDS-C/SR toe te voegen aan de basisset van meetinstrumenten voor de Stichting Benchmark GGZ. Voor de dagelijkse praktijk werd het gebruik van de IDS reeds in 2004 aanbevolen (Nolen & Dingemans 2004) en uit een grote en nog steeds groeiende hoeveelheid literatuur blijkt dat dit een terecht advies is geweest. Hoewel een volledige Nederlandse validatiestudie nog niet beschikbaar is, blijkt uit zowel de beschikbare Nederlandse gegevens als uit internationaal onderzoek op overtuigende wijze dat IDS-C/SR een vragenlijst is met goede psychometrische eigenschappen die in staat is de ernst en verandering bij de depressieve stoornis (én bij subklinische depressieve symptomen) te meten. Alleen de interne consistentie van de vragen voor het meten van atypische klachten bleek minder optimaal (Schulte-van Maaren e.a. 2013). Om die reden wordt de IDS-C/SR zowel in de klinische praktijk als bij het uitvoeren van onderzoek inmiddels veel gebruikt.

De QIDS is een verkorte versie van de IDS en leent zich door zijn korte afnameduur en de overlap met de oorspronkelijke IDS waarschijnlijk nog beter voor toepassing in ROM. Het zonder kosten gebruik kunnen maken van beide instrumenten is een bijkomend, maar niet onbelangrijk, voordeel. Voor de BDI moet namelijk wél steeds betaald worden.

Naast deze wetenschappelijke argumenten is er de toepassing in de praktijk. In tegenstelling tot een weinig passend generiek instrument (zoals de OQ45), zal een behandelaar in de praktijk gebruik willen maken van een meetinstrument dat past bij de specifieke doelgroep (in dit geval de IDS bij depressies) omdat het nut ervan in de behandeling van de patiënt evident is. Als nut en noodzaak samenkomen, is te verwachten dat dit een positief effect zal hebben op het resultaat van de behandeling, en gebruiken we ROM waarvoor deze bedoeld is: het monitoren van de behandeluitkomst bij de individuele patiënt.

LITERATUUR

- Bernstein IH, Rush JA, Carmody AW, Trivedi MH. Item response analysis of the inventory of depressive symptomatology. *Neuropsychiatr Dis Treat* 2006; 2: 557-64.
- Bernstein IH, Rush JA, Carmody AW, Trivedi MH. Clinical vs. self-report versions of the quick inventory of depressive symptomatology in a public sector sample. *J Psychiatr Res* 2007; 41: 239-46.
- Bernstein IH, Rush JA, Suppes T, Kyotoku Y, Warden D. The quick inventory of depressive symptomatology (clinician and self-report versions) in patients with bipolar disorder. *CNS Spectr* 2010; 15: 367-73.
- Biggs MM, Shores-Wilson K, Rush AJ, Carmody TJ, Trivedi MH, Crismon ML, e.a. A comparison of alternative assessments of depressive symptom severity: a pilot study. *Psychiatr Res* 2000; 96: 269-79.
- Blijd-Hoogewys E, Dyck van R, Emmelkamp P, Mulder N, Oude Voshaar R, Schippers G, e.a. Benchmarken is 'werk-in-uitvoering'. *Tijdschr Psychiatr* 2012; 54: 1031-8.
- Brunoni AR, Benute GRG, Fraguas R, Santos NO, Francisco RPV, de Lucia MCS, e.a. The self-rated Inventory of Depressive Symptomatology for screening prenatal depression. *Int J Gynecol Obstetr* 2013; 121: 243-6.
- Corruble E, Legrand JM, Duret C, Charles G, Guelfi JD. IDS-C and IDS-SR: Psychometric properties in depressed in-patients. *J Affect Disord* 1999; 56: 95-101.
- Jong K de, Nugter AM, Lambert MJ, Burlingame GM. Handleiding voor afname en scoring van de outcome questionnaire Q-45.2. Salt Lake City: OQ Measures LLC; 2009.
- Doraiswamy PM, Bernstein IH, Rush AJ, Kyotoku Y, Carmody TJ, Macleod L, e.a. Diagnostic utility of the Quick Inventory of Depressive Symptomatology (QIDS-C₁₆ and QID-SR₁₆) in the elderly. *Acta Psychiatr Scand* 2010; 122: 226-34.
- Egberink IJL, Vermeulen CSM, Frima RM. COTAN Documentatie. Amsterdam: Boom test uitgevers; 2009-2014. (www.cotandocumentatie.nl)
- González DA, Boals A, Jenkins SR, Schuler ER, Taylor D. Psychometrics and latent structure of the IDS and QIDS with young adult students. *J Affect Disord* 2013; 149: 217-20.
- González DA, Jenkins SR. Cross-measure equivalence and communicability in the assessment of depression: a focus on factor-based scales. *Assessment* 2014; 21: 731-41.
- Graaf R de, ten Have M, Tuithof M, van Dorselaer S. Incidentie van psychische aandoeningen. Opzet en eerste resultaten van de tweede meting van de studie NEMESIS-2. Utrecht: Trimbos-Instituut; 2012.
- Hafkenscheid A, van Os J. Huidige ROM doet afbreuk aan valide kwaliteitsmeting. *Tijdschr Psychiatr* 2013; 55: 179-81.
- Hafkenscheid A, van Os J. Naar een deugdelijke ROM. *MGV* 2014; 69: 20-8.
- Hamilton M. Development of a rating scale for primary depressive illness. *Br J Soc Clin Psychol* 1967; 6: 278-96.
- Karsten J, Hartman CA, Ormel J, Nolen WA, Penninx BWJH. Subthreshold depression bases on functional impairment better defined by symptom severity than by number of DSM-IV symptoms. *J Affect Disord* 2010; 123: 230-7.
- Meininger K. Assessment of seasonal affective disorder using self-ratings. A comparison between the IDS-SR, BDI-II and SIGH-SAD-SR. Masterthese Klinische en Ontwikkelingspsychologie. Interne publicatie Rijksuniversiteit Groningen; 2007.
- Nolen WA, Dingemans PMAJ. Meetinstrumenten bij stemmingsstoornissen. *Tijdschr Psychiatr* 46, 10, 681-686.
- Os J van, Kahn R, Denys D, Schoevers RA, Beekman AT, Hoogendijk WJ, e.a. ROM: Gedragsnorm of dwangmaatregel? Commentaar op het themanummer ROM. *Tijdschr Psychiatr* 2012; 54: 245-53.
- Rush AJ, Giles DE, Schlessner MA, Fulton CL, Weissenburger JW, Burns C. The inventory of Depressive Symptomatology (IDS): preliminary findings. *Psychiatry Res* 1986; 18, 65-87.
- Rush AJ, Gullion CM, Baco MR, Jarrett RB, Trivedi MH. The inventory of Depressive Symptomatology (IDS): psychometric properties. *Psychol Med* 1996; 26: 477-86.
- Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, e.a. The 16-item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. *Biol Psychiatry* 2003; 54: 573-83.
- Rush AJ, Trivedi MH, Carmody TJ, Ibrahim HM, Markowitz JC, Keitner GI, e.a. Self-reported depressive symptom measures: sensitivity to detecting change in a randomized, controlled trial of chronically depressed, nonpsychotic outpatients. *Neuropsychopharmacology* 2005; 30: 405-16.
- Schulte-van Maaren YWM, Giltay EJ, Van Hemert AM, Zitman FG, de Waal MWM, Carlier IVE. Reference values for major depression questionnaires: The Leiden Routine Outcome Monitoring Study. *J Affect Disord* 2013; 149: 342-9.
- Sherwood Brown E, Murray M, Carmody TJ, Kennard BD, Hughes CW, Khan DA, e.a. The Quick Inventory of Depressive Symptomatology-Self-report: a psychometric evaluation in patients with asthma and major depressive disorder. *Ann Allergy Asthma Immunol* 2008; 100: 433-8.
- Spitzer RL, Endicott J, Robins E. Research Diagnostic Criteria: Rationale and reliability. *Arch Gen Psychiatry* 1978; 35: 773-82.

- Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. *Psychol Med* 2004; 34: 73-82.
- Wanders RBK, Wardenaar KJ, Kessler RC, Penninx BWJH, Meijer RR, de Jonge P. Differential reporting of depressive symptoms across distinct clinical subpopulations: What difference does it make? *J Psychosom Res* 2015; 78: 130-6.
- Wardenaar KJ, Giltay EJ, van Veen T, Zitman FG, Penninx WJH. Dimensions of the Inventory of Depressive Symptomatology as predictors of the course of depressive and anxiety disorders. *J Psychiatr Res* 2012; 46: 1655-61.

SUMMARY

Reasons for including the Inventory of Depressive Symptomatology (IDS) and its shortened version (QIDS) in routine outcome monitoring by the Foundation Benchmark Mental Health (SBG)

Y. MEESTERS, W.B. DUIJZER, W.A. NOLEN, R.A. SCHOEVERS, H.G. RUHÉ

BACKGROUND Increasingly, health insurance companies are asking for treatment results in mental health care to be expressed in actual numbers. They base their request on the data supplied by SBG (Stichting Benchmark GGZ) and the procedures that SBG has followed. For the purpose of benchmarking, the SBG has accepted a limited number of questionnaires. These tools are not suitable for performing routine outcome monitoring (ROM) in clinical practice. There is a urgent need for these questionnaires to be replaced by a widely-used depression questionnaire, namely the IDS.

AIM To present arguments supporting the view that the IDS is an excellent and useful instrument that the SBG should accept as a measuring tool.

METHOD We discuss the quality of the IDS by reviewing the literature.

RESULTS The IDS is a high quality instrument well suited for measuring the severity of depressions and also is sufficiently sensitive for measuring mood improvements.

CONCLUSION The IDS is an excellent tool for performing ROM measurements and is available free of charge. The authors argue that the SBG should accept this questionnaire as a suitable instrument for benchmarking.

TIJDSCHRIFT VOOR PSYCHIATRIE 58(2016)1, 48-54

KEY WORDS benchmarking, IDS, ROM, SBG