

Behandelresultaten van vrijgevestigde psychologen en psychotherapeuten

Kim de Jong

Tijdschrift voor Psychotherapie

ISSN 0165-1188

Volume 42

Number 5

Tijdschr Psychother (2016) 42:308-317

DOI 10.1007/s12485-016-0148-6



Your article is protected by copyright and all rights are held exclusively by Stichting Tijdschrift voor Psychotherapie. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

Behandelresultaten van vrijgevestigde psychologen en psychotherapeuten

Kim de Jong

Samenvatting

Hoewel er inmiddels diverse artikelen zijn verschenen met behandelresultaten in ggz-instellingen in diverse landen, is er nog weinig bekend over de behandelresultaten bij vrijgevestigden. In dit artikel worden behandelresultaten geanalyseerd uit een grote dataset van vrijgevestigde psychologen en psychotherapeuten. Tevens is gekeken of er voorspellers voor het behandelbeloop gevonden konden worden. Er werden gegevens geanalyseerd van 11.385 cliënten met een begin- en eindmeting op de Outcome Questionnaire (OQ-45) en 1.740 cliënten met ten minste twee metingen op de Outcome Rating Scale (ORS). De resultaten laten zien dat het effect van de behandelingen gemiddeld groot was ($d = 1,12$) en dat ruim twee derde van de cliënten verbeterd of hersteld was na de behandeling. Het hebben van een betaalde baan, arbeidsongeschiktheid, eerdere behandeling en de hoogte van de beginscore waren voorspellend voor het behandelbeloop. Tot slot werd gevonden dat er verschillen waren in de snelheid van verbetering tussen therapieën van verschillende lengte. Deze analyses geven toekomstige cliënten een ijkpunt voor de reële behandelresultaten binnen de vrijgevestigde ggz.

Treatment outcomes in Dutch private practices

Abstract

Although by now several scientific publications on the treatment outcomes in mental health services have appeared, not much is known about treatment outcomes in private practices. This article analyzes the treatment outcomes of a large data set of private practices of psychologists and psychotherapists in the Netherlands. Additionally, potential predictors of treatment outcomes have been

K. de Jong (✉)
Leiden, Nederland
e-mail: k.de.jong@fsw.leidenuniv.nl

assessed. Data from 11,385 clients with a pre and post treatment measurement on the Outcome Questionnaire (OQ-45), as well as from 1,740 clients with at least two measurements on the Outcome Rating Scale (ORS) has been analyzed. The results demonstrated on average a large improvement in functioning ($d = 1.12$) and two third of the clients could be classified as improved or recovered at the end of treatment. Having a paid job, being unfit for work, having had previous treatment, and the severity of distress at intake were predictive of treatment progress. It was also found that there were differences in the rate of change between treatments with different durations. These analyses provide a starting point for future comparisons of treatment outcomes in private practices.

Inleiding

In maart 2016 kwam de Stichting Benchmark GGZ (SBG) naar buiten met de eerste resultaten van de systematische vergelijking (*benchmark*) tussen verschillende ggz-instellingen. Daaruit bleek dat er aanzienlijke verschillen bestaan tussen ggz-instellingen wat betreft behandelresultaten. Tevens kwam naar voren dat één op de drie cliënten met een angst- of stemmingsstoornis geen baat heeft bij behandeling, waarbij 5 % van de cliënten verslechtert en 30 % niet (voldoende) verandert (SBG 2016). Omdat dit de eerste keer is dat resultaten van behandelingen in de Nederlandse praktijk bekend worden gemaakt, is het lastig om te interpreteren of deze resultaten goed of slecht zijn.

Internationaal gezien was de eerste grote naturalistische dataset waarover resultaten gepubliceerd werden afkomstig uit de Verenigde Staten. Hansen et al. (2002) verzamelden resultaten met de *Outcome Questionnaire* (OQ-45; Lambert et al. 2004) bij 6.072 cliënten die in behandeling waren bij bedrijfsartsen en ggz-instellingen. Zij vonden dat 8 % van de cliënten significant verslechterde en 57 % geen (betrouwbare) verandering doormaakte.

In Groot-Brittannië werd (in 2015) een vergelijkbaar onderzoek gedaan, met de *Clinical Outcomes in Routine Evaluation Outcome Measure* (CORE-OM; Evans et al. 2002), die bij ambulante behandelingen in de tweede lijn gebruikt wordt als uitkomstinstrument. Stiles et al. (2015) verzamelden data bij 26.430 patiënten en vonden dat slechts 1 % van de patiënten verslechterde en 19 % geen verandering liet zien. Gegevens over de Britse eerste lijn, van het *Improving Access to Psychotherapy* (IAPT) programma, toonden aan dat daar 6 % verslechtert en 33 % niet verandert (HSCIC 2015).

Er bestaat dus enige variatie tussen de verschillende landen en zorgsystemen. Opvallend daarbij is dat resultaten ook internationaal gezien vrijwel uitsluitend gemeten zijn bij ggz-instellingen, terwijl er over de resultaten van vrijgevestigde therapeuten nog weinig bekend is.

Classificatie van behandelresultaten

Een van de ingewikkeldheden bij het interpreteren van behandelresultaten is dat er verschillende methodes zijn om resultaten te classificeren. Veel wetenschappelijke artikelen gebruiken effectgrootten (Cohens d) om verschillen tussen twee behandelingen uit te drukken. In de jaren tachtig en negentig van de vorige eeuw kwam het gebruik van dit soort maten ter discussie te staan, omdat het een individuele behandelaar weinig informatie geeft over of een behandeling interessant is voor zijn cliënt.

Inmiddels wordt het criterium van klinische significantie (Jacobson en Truax 1991) het meest gebruikt. Deze maat wordt ook in ROM-portals vaak gepresenteerd als behandeluitkomst. De genoemde cijfers over behandelresultaten zijn alle berekend met behulp van deze methode. De methode maakt gebruik van twee criteria om een cliënt te classificeren als hersteld: a een cliënt moet meer verbeteren dan de meetfout van het meetinstrument (betrouwbare verbetering), en b een cliënt moet vergelijkbaar functioneren als mensen die geen behandeling nodig hebben voor psychologische klachten (klinisch significante verbetering). Als er alleen sprake is van betrouwbare verbetering, dan wordt een cliënt geclassificeerd als betrouwbaar verbeterd. Indien een cliënt meer verslechtert is dan de meetfout van het instrument (betrouwbare verslechtering), dan wordt de cliënt geclassificeerd als verslechterd. Als er geen sprake is van betrouwbare verbetering of verslechtering, dan is er sprake van geen (of onvoldoende) verandering.

Een van de kritiekpunten op de methode is dat, om betrouwbare verandering (of verslechtering) te behalen, er zelfs bij een betrouwbare vragenlijst vaak vrij grote veranderingen noodzakelijk zijn, waardoor de groep 'geen verandering' vaak relatief groot is, in vergelijking met andere methodes (Nugter et al. 2013), al hebben andere methodes weer hun nadelen.

ROM-resultaten van vrijgevestigden

In Nederland was het door een combinatie van factoren tot nu toe niet mogelijk voor vrijgevestigde therapeuten om hun ROM-gegevens aan te leveren aan SBG, waardoor de cijfers van SBG dus uitsluitend ggz-instellingen bevat. Wel is het voor vrijgevestigden al enkele jaren verplicht om resultaten te meten bij patiënten.

Een aanzienlijk deel van de vrijgevestigde therapeuten in Nederland is aangesloten bij de Landelijke Vereniging voor Vrijgevestigde Psychologen en Psychotherapeuten (LVVP), voortgekomen uit een fusie van de Nederlandse Vereniging voor Vrijgevestigde Psychotherapeuten (NVVP) en de Landelijke Vereniging Eerstelijnspsychologen (LVE). Deze verenigingen boden leden de mogelijkheid om behandelresultaten te meten, via een online ROM-portal. Inmiddels bevat die portal metingen van verschillende vragenlijsten over ruim zestigduizend cliënten, al heeft lang niet iedere cliënt zowel een begin- als een eindmeting.

In dit artikel worden de gegevens uit de LVVP ROM-portal geanalyseerd, met als doel een eerste indruk te krijgen van de behandelresultaten van vrijgevestigde

Tabel 1 Aantal cliënten met metingen op de door SBG erkende uitkomstinstrumenten.

Vragenlijst	Minimaal één meting	Minimaal twee metingen	Begin- en eindmeting
OQ-45	42.912	21.594	11.385
CORE-OM	7.635	3.700	1.965
SCL-90	2.999	1.361	764
BSI	302	132	74
SQ-48	804	230	146
KKL	476	229	140

psychologen en psychotherapeuten in Nederland. Tevens zal bekeken worden welke achtergrondvariabelen voorspellend zijn voor het behandelbeloop.

Methode

Data

De LVVP ROM-dataset die hier gebruikt is, bevat gegevens die verzameld zijn in de periode januari 2010 tot september 2015. De dataset bevat 290.927 metingen over 62.508 cliënten in 1.115 praktijken. Indien er sprake is van een groepspraktijk, zijn de gegevens doorgaans niet herleidbaar tot een individuele behandelaar.

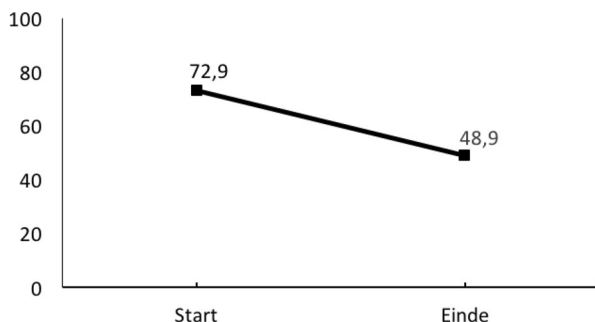
Niet alle data in de portal hebben betrekking op uitkomstmetingen; sommige therapeuten gebruiken de portal ook om diagnostische vragenlijsten af te nemen. In tab. 1 staat vermeld van hoeveel cliënten er gegevens beschikbaar waren in de database, op de door SBG erkende uitkomstinstrumenten voor volwassenen.

Aangezien de resultaten van verschillende meetinstrumenten onderling niet makkelijk vergelijkbaar zijn (Blankers et al. 2016), is er in dit artikel voor gekozen de resultaten van de uitkomstmetingen te beperken tot de OQ-45; veruit het meest-gebruikte meetinstrument in de database. Daarbij is ervoor gekozen uitsluitend cliënten te analyseren van wie bekend is dat er een begin- en eindmeting beschikbaar is.

Naast de door SBG erkende uitkomstinstrumenten, bevat de LVVP-database een grote hoeveelheid metingen met de Outcome Rating Scale (ORS; Miller et al. 2000). Er zijn 1.794 cliënten met ten minste twee metingen op de ORS, die samen 71.359 invullingen hebben. Hierdoor kan van een subgroep cliënten een gedetailleerder beeld verkregen worden van hun vooruitgang in de tijd.

Hoewel de database veel uitkomstdata bevat, is er relatief weinig aanvullende informatie over de cliënten. Er zijn geen achtergrondgegevens bekend over de cliënten in de database, met uitzondering van de cliënten bij wie de aanvullende biografische vragenlijst ($n = 1.172$) is afgenomen. Deze groep wordt apart geanalyseerd in een subgroepanalyse. Ook is niet bekend welke diagnoses de cliënten hadden.

Figuur 1 Behandelvoortgang op de OQ-45 van begin- tot eindmeting.

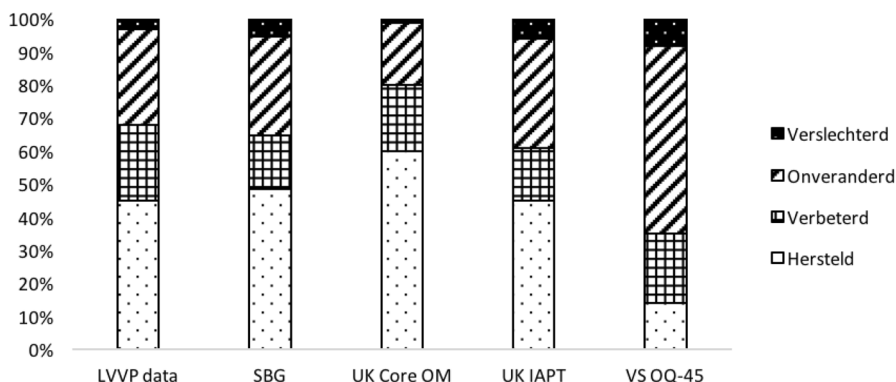


Meetinstrumenten

Outcome Questionnaire (OQ-45). De OQ-45 werd gebruikt om de behandelvoortgang te meten. De OQ-45 (Lambert et al. 2004) is een zelfrapportagevragenlijst en bestaat uit 45 items (waarvan er 9 omgescoord moeten worden) die vragen naar hoe de patiënt zich de afgelopen week heeft gevoeld (0 = nooit; 4 = bijna altijd). De OQ-45 bestaat uit drie subschalen, maar voor dit onderzoek werd alleen de totaalscore gebruikt. De grenswaarde voor normaal functioneren is 55 voor de Nederlandse versie van de OQ-45 en de score voor betrouwbare verandering is 14. De interne consistentie van de Nederlandse OQ-45 varieert van 0,92 tot 0,96 bij groepen studenten, normalen, patiënten en normalen en patiënten gecombineerd (Jong et al. 2009).

Outcome Rating Scale (ORS). De ORS (Miller et al. 2000) werd gebruikt om de behandelvoortgang per sessie te meten. De ORS bestaat uit vier items die elk gemeten worden op een schaal van 10 cm. De items van de ORS meten drie gebieden van functioneren van cliënten: individueel (persoonlijk welbevinden), interpersoonlijk (familie, intieme vrienden) en sociaal (werk, opleiding, sociale contacten). Daarnaast wordt de cliënt gevraagd het algehele welbevinden te scoren. De score van de ORS wordt berekend door de millimeters vanaf rechts op de vier schalen bij elkaar op te tellen. De grenswaarde voor normaal functioneren op de Nederlandse versie van de SRS is 36. De betrouwbaarheid van de Nederlandse ORS is goed, met een interne consistentie van 0,83 (Janse et al. 2014).

Biografische vragenlijst. De biografische vragenlijst is in 2014 vastgesteld door de ROM-commissie van de LVVP en heeft als doel achtergrondgegevens te verzamelen die mogelijk voorspellend zijn voor de behandelresultaten, gebaseerd op wetenschappelijke literatuur en de klinische indruk. De biografische vragenlijst bestaat uit twaalf vragen die gaan over de leeftijd, sekse, etniciteit, leef- en werksituatie, opleidingsniveau, eerdere behandeling, medicatiegebruik bij intake, de hoofdklacht en de duur van de hoofdklacht.



Figuur 2 Klinisch significante verbetering in de LVVP-data en vergelijkingsgroepen.

Resultaten

Behandelresultaten

Er waren 11.385 cliënten met een begin- en eindmeting op de OQ-45 in de database. Deze cliënten werden behandeld in 704 praktijken. De leeftijd van cliënten varieerde van 16-85 jaar, met een gemiddelde van 39,9 jaar ($SD = 13,3$; $n = 9.784$). Het aantal metingen met de OQ-45 per cliënt varieerde van 2-38, waarbij er gemiddeld 2,8 afnames ($SD = 1,8$) waren. Dit duidt erop dat de OQ-45 bij de grootste groep van deze praktijken primair als uitkomstmeting is gebruikt en niet als ondersteunende ROM gedurende de behandeling. De tijd tussen de begin- en eindmeting varieerde van 1-211 weken, met een gemiddelde van 34,8 weken ($SD = 27,3$; $n = 6.679$).

Figuur 1 geeft de verandering op de OQ-45 weer van het begin tot einde van de therapie. Er is flinke variatie in de verandering op de OQ; die loopt van 78 punten verslechterd tot 139 punten verbeterd. Gemiddeld gaan cliënten 24 punten ($SD = 22,2$) vooruit op de OQ-45. Dit komt neer op een effectgrootte van 1,12 (Cohens d), wat gezien wordt als een groot effect.

Als de behandelresultaten geassocieerd worden volgens de criteria van klinische significantie, zijn 5.072 cliënten hersteld (45%), 2.647 cliënten betrouwbaar verbeterd (23%), 3.287 cliënten niet (voldoende) veranderd (29%) en 379 cliënten verslechterd (3%). Figuur 2 vergelijkt de resultaten op basis van de LVVP-data met de bestaande literatuur die in de inleiding is aangehaald. Te zien is dat er in de LVVP-dataset relatief weinig mensen verslechterd zijn, terwijl het percentage cliënten dat hersteld en verbeterd is in het midden zit, vergeleken met de andere datasets.

Tabel 2 Voorspellers voor behandelresultaten.

	B	SE	β
Constance	-16,89	2,74	–
Betaald werk	2,70	1,40	0,06
WAO	-7,60	3,68	-0,06
Eerdere behandeling	-2,58	1,34	-0,06
Score beginmeting	0,57	0,03	0,52

$R^2 = 0,26.$

Voorspellen van behandelresultaten

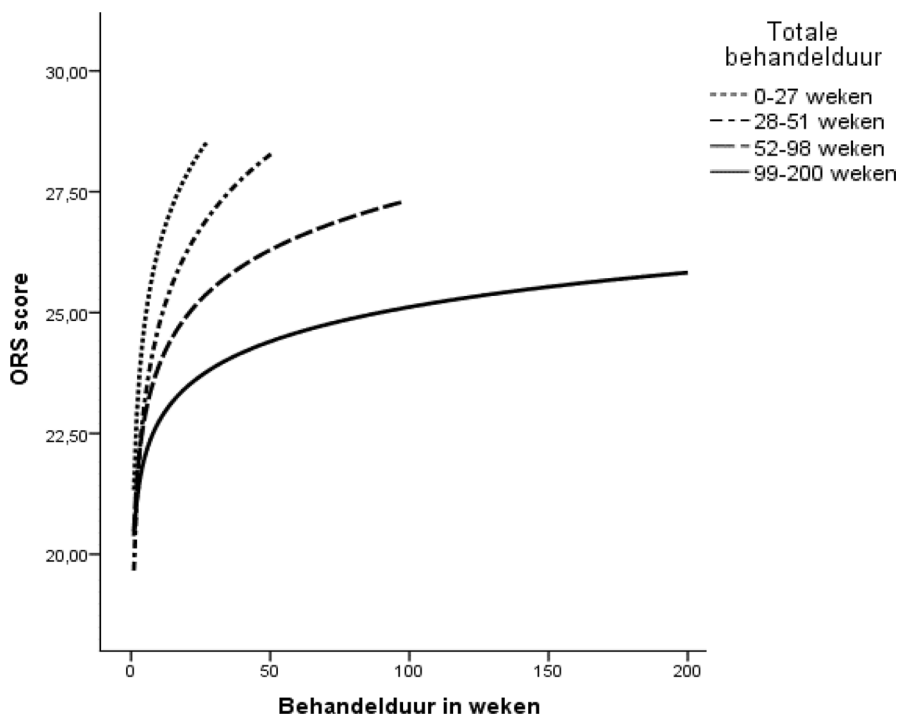
Bij een subgroep van 1.172 cliënten was de biografische vragenlijst aanwezig. Bij deze groep werd gekeken of er achtergrondvariabelen waren die voorspellend waren voor de uitkomst van de behandeling. Er werd gekeken naar het effect van beginscore, leeftijd, etniciteit, werksituatie, woonsituatie, opleidingsniveau en eerdere behandeling. Met behulp van een achterwaartse multiële regressieanalyse werd een model gevonden met vier voorspellers, die samen 26 % van de variantie in de verandering in klachten gedurende de behandeling verklaren. Het hebben van betaald werk en een hogere beginscore voorspellen een grotere verandering op de OQ-45. Cliënten die in de WAO zitten of al eerder behandeling hebben gehad voor psychische problemen, hebben minder positieve prognoses (zie tab. 2).

Behandelvoortgang op de ORS

Er waren 1.794 cliënten met ten minste twee metingen op de ORS. Deze cliënten werden behandeld in 390 praktijken. De leeftijd van de cliënten varieerde van 16–84 jaar, met een gemiddelde van 39,1 jaar ($SD = 13,0$; $n = 1.421$). Het aantal metingen met de ORS per cliënt varieerde van 2–208, waarbij er gemiddeld 10,1 afnames ($SD = 15,5$) waren.

Met behulp van een multilevelanalyse werd getoetst in hoeverre er sprake is van verbetering in scores over tijd op de ORS. Er was sprake van een significant log-lineair patroon, dat gekenmerkt wordt door snelle vooruitgang in de beginfase van de behandeling en langzamere vooruitgang later in de behandeling, $t(6.240) = 35,3$, $p < 0,001$. Er werd gekeken of er op de ORS sprake was van een voorspellende waarde van de beginscore. In tegenstelling tot veel psychotherapieonderzoek bleek de beginscore niet voorspellend te zijn voor de behandelvoortgang op de ORS.

Onderzoek van Baldwin et al. (2009) heeft laten zien dat behandelingen met een verschillende behandelduur doorgaans verschillende patronen laten zien. Daarom werd de behandelduur ingedeeld in vier gelijke groepen. De vier groepen lieten inderdaad een verschillend patroon zien (zie fig. 3). Er werd geen verschil gevonden wat betreft de score aan het begin van de behandeling, maar wel wat betreft de snelheid waarmee klachten verbeterden, waarbij de verbetering bij korte behandelingen sneller ging dan bij lange behandelingen.



Figuur 3 Behandelvoortgang op de ORS over tijd.

Discussie

Dit artikel beschrijft de eerste analyse van behandelresultaten bij vrijgevestigde psychologen en psychotherapeuten in Nederland. Uit de resultaten komt naar voren dat de behandelingen gemiddeld genomen een groot effect hebben en dat – net als bij cliënten met angst- en stemmingsstoornissen in de volwassen curatieve ggz – circa twee derde van de cliënten positieve behandelresultaten heeft (hersteld en verbeterd).

Ook in deze dataset vinden we dat ongeveer een derde van de cliënten niet (voldoende) profiteert van de behandeling. Vergeleken met de gegevens van SBG is in deze dataset het percentage cliënten dat verslechtert gedurende de therapie wel iets lager.

Verder werd gevonden dat een aantal demografische variabelen voorspellend is voor een positiever of negatiever behandelbeloop: cliënten die reeds eerdere behandeling hebben gehad en cliënten die in de WAO zitten gaan minder vooruit gedurende de behandeling, terwijl cliënten met een betaalde baan en cliënten met een hoge beginscore meer vooruitgang boeken. Dit laatste lijkt wellicht niet intuïtief, maar is te verklaren vanuit het feit dat cliënten met een hogere beginscore ook meer ruimte hebben om te verbeteren.

Deze analyses laten, behalve de feitelijke resultaten, ook zien hoe ingewikkeld het is om dit soort resultaten te interpreteren, zelfs al zijn de aantallen groot. Zonder cijfers om de gegevens tegen af te zetten, weten we niet wat goede resultaten zijn en wat niet. Vergelijking met eerder in de literatuur gepubliceerde resultaten biedt wel enige context, maar heeft ook enkele problemen. De dataset van SBG die gepubliceerd is, bevat alleen cliënten met angst- en stemmingsstoornissen. De LVVP-dataset omvat waarschijnlijk een bredere groep cliënten. Binnen de IAPT-data is gemeten met stoornisspecifieke meetinstrumenten, waarvan we weten dat deze gevoeliger zijn voor verandering dan generieke meetinstrumenten, zoals de OQ-45 en CORE-OM (Fletcher et al. 1992).

Toch is het geruststellend dat in deze dataset de resultaten redelijk vergelijkbaar zijn met de resultaten die SBG gepubliceerd heeft en de IAPT-data. Opvallend zijn vooral de slechte scores in de Verenigde Staten en de goede resultaten in de CORE-OM-data. Daarbij is waarschijnlijk ook van belang dat in de Verenigde Staten de toegang tot zorg slechter geregeld is, en dat de opleiding tot therapeut daar ook minder gereguleerd is dan in Nederland. Voor de goede resultaten in de CORE-OM-data is niet zo makkelijk een verklaring te geven.

Deze data-analyse heeft nog een aantal andere beperkingen. Zo is niet duidelijk hoe representatief deze gegevens zijn voor vrijgevestigde behandelaren in Nederland. Het is mogelijk dat LVVP-leden een selectieve groep zijn, bijvoorbeeld doordat ze vaker dan gemiddeld een bepaalde therapeutische oriëntatie hebben. Ook is er heel weinig bekend over de cliënten in deze dataset. Er is maar van een kleine groep achtergrondgegevens bekend, er zijn geen diagnoses en het is niet duidelijk hoeveel zorg een cliënt precies heeft gehad.

Daarnaast heeft ook in de selectie van de gegevens wellicht invloed op de resultaten. Alleen de gegevens van de OQ-45 en de ORS zijn geanalyseerd; mogelijk hadden andere meetinstrumenten (o.a. SCL-90, BSI, CORE-OM) tot andere resultaten geleid. Wellicht gebruiken bepaalde behandelaren bepaalde meetinstrumenten. Ook de keuze om alleen volledige behandeltrajecten te analyseren, heeft mogelijk invloed op de resultaten. Het is goed denkbaar dat niet alle behandelaren netjes registeren wanneer er een eindmeting is gedaan. Zonder gegevens over de DBC is dat echter niet te bepalen. Wanneer alle cliënten met ten minste twee metingen geanalyseerd zouden worden, kwamen de resultaten lager uit, maar dat kan er ook aan liggen dat er dan nog veel niet-afgeronde behandelingen tussen zitten, zeker voor behandelaren die ROM gebruiken als onderdeel van de behandeling.

Kortom, er zijn de nodige beperkingen aan deze data-analyse, maar de gegevens kunnen gebruikt worden als een eerste ijkpunt voor de resultaten van behandelingen bij vrijgevestigde psychologen en psychotherapeuten. Om in de toekomst betere interpretaties te kunnen doen, is het belangrijk dat er relevante achtergrondinformatie beschikbaar is over cliënten, met name over variabelen die mogelijk voorspellend zijn voor de behandeluitkomst. Wanneer behandelaren frequente metingen doen (zoals in deze dataset bij de ORS), krijgen we ook meer inzicht in het beloop van behandelingen. Deze informatie kan in de toekomst gebruikt worden om al tijdens de behandeling de koers eventueel bij te stellen, of om een prognose te geven van de behandelduur.

Literatuur

- Baldwin, S. A., Berkeljon, A., Atkins, D. C., Olsen, J. A., & Nielsen, S. L. (2009). Rates of change in naturalistic psychotherapy: contrasting dose-effect and good-enough level models of change. *Journal of Consulting and Clinical Psychology, 77*, 203–211.
- Blankers, M., Barendregt, M., & Dekker, J. J. M. (2016). Meetvariatie als bron van bias bij het benchmarken met verschillende ROM instrumenten. *Tijdschrift voor Psychiatrie, 58*, 55–60.
- Evans, C., Connell, J., Barkham, M., Margison, F., McGrath, G., Mellor-Clark, J., & Audin, K. (2002). Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *British Journal of Psychiatry, 180*, 51–60.
- Fletcher, A., Gore, S., Jones, D., Fitzpatrick, R., Spiegelhalter, D., & Cox, D. (1992). Quality of life measures in health care, II. Design, analysis, and interpretation. *British Medical Journal, 305*, 1145–1148.
- Hansen, N. B., Lambert, M. J., & Forman, E. M. (2002). The psychotherapy dose-response effect and its implications for treatment delivery services. *Clinical Psychology: Science and Practice, 9*, 329–343.
- Health and Social Care Information Centre (2015). *Psychological therapies: annual report on the use of IAPT services. England, 2014/15*. Leeds: Health and Social Care Information Centre.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12–19.
- Janse, P. D., Boezen-Hilberdink, L., Dijk, M. K. van, Verbraak, M. J. P. M., & Hutschemaekers, G. J. M. (2014). Measuring feedback from clients: the psychometric properties of the dutch outcome rating scale and session rating scale. *European Journal of Psychological Assessment, 30*, 86–92.
- Jong, K. de, Nugter, M. A., Lambert, M. J., & Burlingame, G. M. (2009). *Handleiding voor afname en scoring van de Outcome Questionnaire (OQ-45)*. Salt Lake City: OQ Measures LLC.
- Lambert, M. J., Morton, J. J., Hatfield, D. R., Harmon, C., Hamilton, S., Shimokawa, K., Christopherso, C., Pearl, S., Burlingame, G. M., & Reisinger, C. W. (2004). *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Wilmington: American Professional Credentialing Services LLC.
- Miller, S. D., Duncan, B. L., & Johnson, L. (2000). *Outcome rating scale*. Chicago: Institute for the Study of Therapeutic Change.
- Nugter, M. A., Jong, K. de, & Kortrijk, H. E. (2013). Meten van veranderingen bij patiënten. In V. J. A. Buwalda, M. A. Nugter, W. van Tilburg & A. T. F. Beekman (red.), *Praktijkboek ROM in de ggz II. Implementatie en gebruik bij verschillende doelgroepen* (pag. 149–164). Utrecht: de Tijdstroom.
- Stichting Benchmark GGZ (2016). *Praktijkvariatie in uitkomst GGZ behandelingen*. www.zorgvisie.nl/PageFiles/92627/Stichting%20Benchmark%20GGZ%20bijlage.pdf. Geraadpleegd op: 18. Juli 2016.
- Stiles, W. B., Barkham, M., & Wheeler, S. (2015). Duration of psychological therapy: relation to recovery and improvement rates in UK routine practice. *British Journal of Psychiatry, 207*, 115–122.

Dr. K. de Jong is werkzaam als universitair docent bij de Universiteit Leiden en als basispsycholoog bij Psychologie La Croix te Amsterdam. Zij is tevens bestuurslid op de portefeuille Kennis en ROM bij de VGCT.