

Behandeluitkomsten gespecialiseerde ggz-behandelaren: ggz-instelling versus vrijgevestigd

Rudy Douven, Tobias Hlobil en Lianne Schippers

Patiënten in de specialistische ggz kunnen behandeld worden in een ggz-instelling of door vrijgevestigde behandelaren. We onderzoeken voor initiële behandelingen in 2008-2014 of de kans op een toename van het algemeen functioneren van een patiënt, zoals gemeten door de GAF-score, verschilt tussen beide type behandelaren. De GAF-score is een score die behandelaren zelf rapporteren voor en na een behandeling. We vinden dat de kans op een toename van de GAF-score, voor patiënten met dezelfde diagnose en mate van klachten, ongeveer anderhalf keer hoger is bij vrijgevestigde behandelaren dan bij een ggz-instelling. In dit artikel bespreken we mogelijke verklaringen voor deze verschillen.

Trefwoorden: specialistische ggz, ggz-instellingen, vrijgevestigde behandelaren, GAF-score.

Inleiding

Jaarlijks behandelt de specialistische geestelijke gezondheidszorg (s-ggz) ongeveer een half miljoen mensen (NZa, 2019). Deze mensen ontvangen behandelingen voor zwaardere of complexere psychische klachten. De overgrote meerderheid van deze behandelingen (86%) vindt plaats in ggz-instellingen door behandelaren in loondienst. Een kleine, maar groeiende minderheid van de s-ggz behandelingen (14%) wordt geleverd door vrijgevestigde behandelaren (Vektis, 2017). Dit roept de vraag op of dit een verschil maakt voor de uitkomst van de behandeling. Heeft een patiënt meer kans om vooruit te gaan bij behandelaren in dienst bij ggz-instellingen, of bij vrijgevestigde behandelaren?

Een reden om een verschil in behandeluitkomst te verwachten is dat de financiële prikkels tussen behandelaren verschillen. Vrijgevestigde behandelaren kregen tot vrij recent betaald per diagnose-behandelcombinatie (DBC). Ggz-instellingen daarentegen waren gebonden aan een budget en de behandelaren ontvingen een vast bedrag als salaris. Dit betekent dat vrijgevestigde behandelaren en behandelaren in dienst bij ggz-instellingen andere financiële prikkels hadden. En recent onderzoek laat zien dat behandelaren en ggz-instellingen niet ongevoelig zijn voor financiële prikkels (zie Douven et al., 2015; Douven et al., 2021). Dit is in lijn met een bredere literatuur in de gezondheidseconomie dat financiële prikkels gevolgen kunnen hebben voor behandeluitkomsten (Chandra et al., 2012).

Een andere reden om een verschil in behandeluitkomsten te verwachten is vanwege zelfselectie. Het verschil in financiële prikkels en andere arbeidsomstandigheden zijn voor sommige mensen reden om te kiezen voor het werk in een ggz-instellingen, en voor andere

mensen juist reden om te kiezen voor werk als een vrijgevestigde behandelaar. Joldersma (2019) beschrijft de specifieke pull en push-factoren om te werken voor een ggz-instellingen. Pull-factoren om voor een ggz-instelling te werken zijn werken in multidisciplinaire teams, hart voor de organisatie, baanzekerheid, en ontwikkelingsmogelijkheden. Push-factoren om niet voor een ggz-instelling te werken zijn werkdruk, autonomie op het werk, en het salaris. Deze pull en push-factoren worden door behandelaren anders gewaardeerd, waardoor behandelaren die kiezen om te werken voor een ggz-instelling verschillen van behandelaren die kiezen om vrijgevestigd te zijn. Deze zelfselectie kan tot gevolg hebben dat ook de kwaliteit van een behandeling verschilt tussen een ggz-instelling en vrijgevestigde behandelaren (Joldersma, 2019).

Hoewel een verschil in behandeluitkomsten in theorie niet onaannemelijk is, is niet eerder onderzocht of deze verschillen ook optreden in de praktijk. Dit onderzoek draagt bij aan de literatuur door de toename van General Assessment of Functioning scores (hierna: GAF-scores) van patiënten behandeld door ggz-instellingen te vergelijken met patiënten behandeld door vrijgevestigde behandelaren. De GAF-score is een eendimensionale maat waarmee behandelaren het psychisch, sociaal en beroepsmatig functioneren van een persoon scoren. Behandelaren in de s-ggz waren jarenlang verplicht om de GAF-score te rapporteren bij aanvang en na afloop van een behandeling. Dit biedt ons de gelegenheid om de toename van GAF-scores tussen verschillende type aanbieders te vergelijken.

Voor deze vergelijking is het belangrijk om rekening te houden met mogelijke verschillen in de patiëntpopulatie tussen ggz-instellingen en vrijgevestigde behandelaren. Ggz-instellingen bieden namelijk een breder palet van zorg aan dan vrijgevestigde behandelaren. Hierbij kan gedacht worden aan crisiszorg, bemoeizorg en een opname. Dergelijke zorg mag niet door vrijgevestigde behandelaren aangeboden worden. Eerder onderzoek vindt bijvoorbeeld wel dat na behandelingen door vrijgevestigde behandelaren de GAF-score meer toeneemt dan na behandelingen door ggz-instellingen, maar houdt geen rekening met populatieverschillen tussen beide typen behandelaren (Zoutenbier en Douven, 2016).

De onderzoeksvraag die dit artikel wil beantwoorden luidt:

“Is de kans voor de patiënt op een toename van de GAF-score bij vrijgevestigde behandelaren gelijk aan die bij behandelaren in dienst bij een ggz-instelling voor patiënten met dezelfde (sub-)diagnose en mate van klachten?”

Voor het beantwoorden van deze onderzoeksvraag wordt hieronder eerst het gebruik van de GAF-score als uitkomstmaat verder toegelicht. Vervolgens worden de onderzoekspopulatie en de onderzoeksmethoden beschreven, gevolgd door de resultaten. Tot slot volgt een discussie van de resultaten en een conclusie met ons antwoord op de onderzoeksvraag.

GAF-score

Het meten van de kwaliteit van een behandeling in de geestelijke gezondheid is complex. Kwaliteit heeft verschillende aspecten zoals kwaliteit van leven, cliëntervaring, functioneren van cliënt, herstel, of risico op terugval (Jongebreur et al., 2015). In de praktijk meten de beschikbare kwaliteitsindicatoren, zoals ROM ('Routine Outcome Measuring), PROMS ('Patient Reported Outcome Measures'), HONOS ('Health of the Nation Outcomes Scale'), en de GAF-score, slechts een zeer beperkt aantal aspecten van kwaliteit.

In dit onderzoek hebben we de beschikking over één indicator, namelijk de GAF-score. De GAF-score maakt onderdeel uit van DSM-IV en behandelaren in de s-ggz scoren patiënten op de GAF-schaal van 1 tot 10 bij het openen van een DBC en bij het sluiten van een DBC.¹ De GAF-score zegt iets over het functioneren van patiënten; mensen zonder significante psychische, beroepsmatige of sociale problemen hebben een score van ongeveer 9 tot 10 en voor mensen met (meer van) deze problemen zal de score lager zijn.

De betrouwbaarheid van de GAF-score als uitkomstmaat heeft beperkingen. In de literatuur bestaan zorgen over het gebruik van de GAF-score om patiënten onderling te vergelijken, omdat verschillende behandelaren een andere GAF-score kunnen toewijzen aan dezelfde patiënt (Grootenboer et al., 2012). De GAF-score kan daarom niet gebruikt worden om individuele patiënten te vergelijken, maar wel om groepen patiënten met elkaar te vergelijken (Soderberg et al., 2005; Urbanoski et al., 2014). Door gemiddelden van groepen te vergelijken wordt de willekeurige variatie die optreedt bij het scoren van individuele patiënten immers verminderd.

De validiteit van de GAF-score als uitkomstmaat heeft ook beperkingen. En wel om dezelfde reden: behandelaren kunnen dezelfde patiënt een andere GAF-score toekennen. Maar dit kan ook een strategische reden hebben (Greenberg en Rosenbeck, 2005). Een voorbeeld daarvan is wanneer de GAF-score wordt gebruikt door externe instanties om de gezondheid van een populatie te beoordelen. In dat geval kan de instelling belang hebben bij het anders invullen van de GAF-score en ontstaan er problemen met de validiteit. Greenberg en Rosenbeck (2005) vinden in hun dataset soms grote variaties in GAF-scores die ze verklaren door een verschil in het administreren van GAF-scores. In de discussie komen we op deze argumenten terug.

Uit eerder onderzoek met de GAF als uitkomstmaat is het volgende bekend. Bijna de helft van de patiënten in de Nederlandse ggz gaat na behandeling vooruit met één GAF-punt (Verhaak et al., 2013).² Wel gaat dit over patiënten in de eerste lijn, die vaker lichtere psychische klachten hebben. En de zwaarte van de psychische klachten is een belangrijke

¹ In de DSM-IV is de GAF-score een waarde tussen de 0-100. Echter, bij het invullen moesten zorgaanbieders een 'schatting' geven per 10 punten, daarom observeren we in onze data een schaal van 0-10.

² In Verhaak et al. (2013) is een schaal van 1 tot 100 gebruikt in plaats van een schaal van 1 tot 10. De patiënten gingen dus 10 punten of meer vooruit. Dit was op basis van een steekproef van 55.067 patiënten behandeld door 613 eerstelijnszorg psychologen.

indicator voor of de GAF-score toeneemt na behandeling (Warmerdam et al., 2016; Söderberg en Tungström, 2007). Ook hangt de toename in GAF-score af van de diagnose aan het begin van een behandeling, het geslacht, het opleidingsniveau, en het al dan niet hebben van een migratieachtergrond (Verhaak et al., 2013; Sonesson et al., 2013). Een onderzoek tussen vrijgevestigde behandelaren laat zien dat de GAF-score voor vergelijkbare patiënten minder toeneemt wanneer die behandeld worden door financieel gemotiveerde behandelaren, i.e. behandelaren die strategisch gebruikmaken van de betalingsstructuur om hun inkomen te vergroten, dan door altruïstische behandelaren, i.e. behandelaren die geen gebruikmaken van de betalingsstructuur (Douven et al., 2017). Tot slot, laat onderzoek zien dat sommige ggz-instellingen patiënten langer gingen behandelen na een negatieve vraagschok, maar dat dit langer behandelen niet leidde tot een toename in de GAF-score (Douven et al., 2021).

Onderzoekspopulatie

Voor het vergelijken van de kans op een toename in de GAF-score na behandeling door ggz-instellingen en vrijgevestigde behandelaren wordt gebruikgemaakt van een administratieve dataset van de Nederlandse Zorgautoriteit. Deze dataset bestaat uit alle DBC's in de s-ggz in Nederland over de periode 2008-2014. Voor elke DBC bevat deze dataset informatie over zowel de patiënt als de behandelaar.

Om alternatieve verklaringen uit te sluiten zijn DBC's geselecteerd die naar verwachting vergelijkbaar zijn voor ggz-instellingen en vrijgevestigde behandelaren. We kiezen daarom alleen DBC's die volledig zijn ingevuld, op een reguliere manier zijn afgesloten door de behandelaar, en waar sprake was van een directe behandeltijd van ten minste 100 minuten. Van deze DBC's kan gezegd worden dat er daadwerkelijk een behandeling heeft plaatsgevonden. Een andere keuze is om alleen DBC's mee te nemen met een behandeltijd van minder dan 6000 minuten, omdat deze nauwelijks voorkomen bij vrijgevestigde aanbieders. Ook DBC's die niet voorkomen bij vrijgevestigde behandelaren, zoals crisiszorg of DBC's met verblijf, nemen we niet mee in de analyse. Tot slot beschouwen we alleen unieke patiënten, patiënten die één keer in dataset voorkomen en waarvoor geen vervolgd-DBC's bestaan.³

Deze selectie heeft geleid tot een dataset met de volgende beschrijvende statistieken (zie tabel 1). De populatie bestaat uit circa 700.000 behandelingen of DBC's: 105.570 behandelingen door vrijgevestigde behandelaren en 593.800 behandelingen door ggz-instellingen. Deze behandelingen zijn uitgevoerd door 1739 vrijgevestigde behandelaren en 893 ggz-instellingen. De gemiddelde leeftijd van de patiënten is 38 jaar. Vrijgevestigde behandelaren behandelen relatief meer vrouwen, en stellen minder verschillende sub-

³ Patiënten met een vervolgd-DBC komen verhoudingsgewijs veel vaker voor bij instellingen. Zoutenbier en Douven (2016) laten bovendien zien dat verbeteringen in de GAF-score gemiddeld genomen kleiner is bij vervolghandelingen. Voor patiënten uit 2008 (of 2014) kunnen we niet nagaan of ze in het voorgaande (of daaropvolgende) jaar een behandeling hebben gehad. Daarom beschouwen we in een robuustheidsanalyse ook alleen patiënten met een behandeling in 2009-2013.

diagnoses. Ook is de geografische spreiding van patiënten iets kleiner voor vrijgevestigde behandelaren.

Tabel 1 Beschrijvende statistiek van de onderzoekspopulatie (2008-2014)

	Vrijgevestigd	Instellingen
Aantal unieke patiënten	105.570	593.800
Aantal aanbieders	1.739	893
Gestelde sub-diagnoses (% t.o.v. alle 723 sub-diagnoses)	59,5	97,0
Postcodes van patiënten (% t.o.v. alle 914 postcodes*)	92,0	98,9
Gemiddelde leeftijd patiënt (in jaren)	37,5 (16,0)	38,2 (21,9)
Vrouwen (%)	60,8 (48,8)	53,9 (49,8)
Gemiddelde totale behandelduur (minuten)	952,2 (749,6)	976,5 (762,6)
Gemiddelde directe behandelduur (minuten)	641,6 (464,8)	557,1 (451,6)
Gemiddelde indirecte behandelduur (minuten)	310,6 (274,5)	419,4 (400,5)
Gemiddelde GAF-score bij aanvang van behandeling	6,22 (0,83)	5,96 (1,01)
Proportie patiënten met vooruitgang (%)	79,5 (40,4)	44,1 (49,6)

Noot: Standaardfouten tussen haakjes.

* Het betreft hier driecijferige postcodegebieden.

De gemiddelde totale behandelduur is nagenoeg hetzelfde bij vrijgevestigde behandelaren en ggz-instellingen, maar vrijgevestigde behandelaren hebben gemiddeld meer direct contact met patiënten dan behandelaren in dienst van ggz-instellingen. Dat ggz-instellingen gemiddeld meer indirecte tijd besteden dan vrijgevestigde behandelaren, zoals tijd voor administratie, onderlinge afstemming en coördinatie, komt overeen met het verschil in werkwijze, namelijk dat instellingen vaker met multidisciplinaire teams werken en vrijgevestigde behandelaren vaker zelf de behandeling uitvoeren.

De toename van de GAF-score is bij een vrijgevestigde aanbieder bijna twee keer zo groot als bij een ggz-instelling: de verhouding patiënten die vooruitgaan bij een vrijgevestigde behandelaar is 80% en slechts 44% bij een ggz-instelling. Hierbij is echter nog geen rekening gehouden met patiëntkenmerken. Bijvoorbeeld, een mogelijke verklaring die uit tabel 1 volgt is dat de gemiddelde patiëntzwaarte bij aanvang van een behandeling, gemeten in GAF-score, met 5,96 iets lager is bij een ggz-instelling dan 6,22 bij vrijgevestigde behandelaren.

Onderzoeksmethode

Het verschil in behandeluitkomst tussen een ggz-instelling en vrijgevestigde behandelaren wordt gemeten aan de hand van de GAF-score. Om de vooruitgang van een behandeling voor patiënt i te meten is de indicator variabele $ToenameGAF_i$ als volgt gedefinieerd.

$$ToenameGAF_i = \begin{cases} 1, & eindGAF_i > startGAF_i \\ 0, & eindGAF_i \leq startGAF_i \end{cases}$$

In deze formule is $startGAF_i$ en $eindGAF_i$ de GAF-score die patiënt i krijgt aan het begin en aan het einde van een behandeling. Een waarde van 1 geeft aan dat de patiënt vooruit is gegaan na behandeling en een waarde van 0 geeft aan dat de patiënt niet vooruit is gegaan. We kiezen voor een binaire afhankelijke variabele omdat het overgrote deel van de behandelingen een verschil in GAF-score hebben van 0 punten (49%), 1 punt (27%) en 2 punten (16%) kent. Een vermindering vindt slechts plaats in 2% van de behandelingen, en een vooruitgang van meer dan 2 punten in 6% van de behandelingen.

Om zo goed mogelijk te corrigeren voor patiëntkenmerken schatten we een lineair regressiemodel. We kiezen voor een OLS regressie in plaats van een probit of logistische regressie omdat bij een binaire afhankelijke variabele, en voldoende grote datasets, het verschil tussen een probit of logistische en OLS-benadering klein is (Angrist en Pischke, 2008).

$$\begin{aligned} ToenameGAF_i = & \beta_0 + \beta_1 type\ aanbieder_i + \beta_2 leeftijd_i + \beta_3 geslacht_i + \beta_4 startGAF_i \\ & + \beta_5 diagnose_i + \beta_6 postcode_i + \beta_7 jaar_i + \beta_8 sluitingsreden_i \\ & + \beta_9 behandelduur_i + \varepsilon_i \end{aligned}$$

De nulhypothese die wordt getoetst is dat $\beta_1 = 0$ waarbij $type\ aanbieder_i$ een indicator variabele is met 1 wanneer sprake is van een behandeling door een vrijgevestigde aanbieder en 0 wanneer sprake is van een behandeling door een ggz-instelling. De interpretatie van β_1 is de kans op toename in de GAF-score in procentpunt.

Met deze specificatie wordt rekening gehouden met een aantal patiëntkenmerken. Om rekening te houden met het feit dat vrijgevestigde aanbieders voornamelijk patiënten met lichtere psychische klachten behandelen dan een ggz-instelling, wordt de GAF-score (9 indicatoren) en de subdiagnose (722 indicatoren) bij aanvang van een behandeling meegenomen. Andere patiëntkenmerken die de kans op vooruitgang kunnen verhogen of verlagen zijn geslacht, leeftijd, opleidingsniveau, inkomen, en migratieachtergrond. Voor geslacht (1 indicator) en leeftijd (18 indicatoren) is data beschikbaar om daar direct voor te controleren. Voor opleidingsniveau, inkomen, en migratieachtergrond, alsook andere kenmerken, kan niet direct worden gecontroleerd. Driecijferige postcodes (913 indicatoren) worden gebruikt als een proxy hiervoor. Een langere behandelduur (17 indicatoren) en reden van sluiten (2 indicatoren) zijn mogelijk ook een reden voor patiëntverschillen.⁴ Een kanttekening hierbij is dat deze twee variabelen ook een direct effect hebben op de toename in de GAF-score, en meten we de behandeluitkomst conditioneel op de lengte en reden van sluiten van een behandeling. Wanneer we echter alleen geïnteresseerd zijn in het verschil in behandeluitkomst tussen beide type behandelaren dan is deze correctie niet nodig. Daarom zullen we ook de regressieresultaten laten zien zonder deze twee variabelen. Als laatste houden we rekening met het jaar waarin

⁴ De drie mogelijke sluitredenen zijn: behandelaar sluit behandeling om inhoudelijke redenen, behandelaar en patiënt eindigen behandeling in overleg of de behandelaar sluit behandeling om administratieve redenen.

de behandeling is gestart (6 indicatoren), om rekening te houden met jaar-effecten, zoals hervormingen in het beleid. De storingstermen clusteren we op het niveau van de aanbieder om rekening te houden met de mogelijkheid dat de GAF-scores van patiënten binnen een aanbieder gecorreleerd zijn.

Resultaten

De regressieresultaten in model (1) van tabel 2a laten zien dat de kans dat een patiënt na behandeling vooruitgaat 27,4 procentpunt hoger is bij vrijgevestigde behandelaren dan bij ggz-instellingen. De kans op een waarde gelijk aan of groter dan 27,4 procentpunt onder de nulhypothese is uitermate klein gegeven de standaardfout van 2,7 procentpunt. Dit suggereert dat de kans op een toename van de GAF-score bij een patiënt meer dan anderhalf keer hoger is bij een vrijgevestigde aanbieder dan bij een ggz-instelling.⁵

Tabel 2a Resultaten en Robuustheidsanalyse 1: Specificatie controlevariabelen

Afhankelijke variabele: vooruitgang (1=GAF-score eind behandeling > GAF-score start behandeling; 0 anderszins)			
	Vershil in vooruitgang (procentpunt)	R ²	Aantal observaties
(1) Model met alle controle variabelen	27,4*** (2,7)	0,237	699.370
(2) Model zonder controle variabelen	35,3*** (3,1)	0,064	699.370
(3) Als (1), zonder behandelduur en sluitreden	30,3*** (2,9)	0,183	699.370

Standaardfouten tussen haakjes zijn geclusterd op aanbiedersniveau. Gebaseerd op een tweezijdige test geven *, ** en *** respectievelijk de significantie op 0,10, 0,05 en 0,01 niveau weer.

Model (2) is het model zonder controlevariabelen. Wanneer we dus geen rekening houden met patiëntkarakteristieken is de geschatte kans op een toename in de GAF-score na behandeling 35,3 procentpunt hoger bij vrijgevestigde behandelaren dan bij ggz-instellingen. De controlevariabelen in de regressievergelijking kunnen dus maar 35,3-27,4=7,9 procentpunt van het verschil verklaren. In model (3) worden, zoals hierboven aangegeven, de behandelduur en de reden voor sluiting van een DBC weggelaten. De geschatte kans is hier 30,3 procentpunt hoger bij vrijgevestigde behandelaren. Ook hier is de kans op een toename in de GAF-score significant groter bij vrijgevestigde behandelaren dan voor een ggz-instelling.

Tabel 2b laat de regressieresultaten van enkele robuustheidsanalyses zien waarbij we weer alle controlevariabelen meenemen. In model (a) laten we de jaren 2008 en 2014 weg uit de analyse omdat dit de kans vergroot dat alle patiënten die we meenemen precies één

⁵ Uit de laatste rij in tabel 1 volgt dat de kans op verbetering 44,1 procentpunt is bij ggz-instellingen. Na correctie voor patiëntkenmerken volgt dat deze kans 27,4 procentpunt hoger is voor vrijgevestigde behandelaren. De toename in GAF score is dus meer dan anderhalf keer zo hoog $((44,1+27,4)/44,1= 1,6)$ bij vrijgevestigde behandelaren.

behandeling hebben gehad. Ook hier vinden we een 26,6 procentpunt hogere kans voor toename bij patiënten onder behandeling van vrijgevestigde behandelaren.

Tabel 2b Robuustheidsanalyse 2: Subset onderzoekspopulatie

Afhankelijke variabele: vooruitgang (1=GAF-score eind behandeling > GAF-score start behandeling; 0 anderszins)			
	Vershil in vooruitgang (procentpunt)	R ²	Aantal observaties
(a) Zonder de jaren 2008 en 2014	26,6*** (2,8)	0,238	522.876
(b) Alleen diagnose ontwikkelingsstoornis	18,8*** (3,9)	0,154	109.172
(c) Alleen diagnose persoonlijkheidsstoornis	33,2*** (3,9)	0,297	28.544
(d) Alleen behandelingen met start GAF=6	26,8*** (2,7)	0,244	329.475
(e) Alleen behandelingen met start GAF=7	26,8*** (3,5)	0,211	157.737

Standaardfouten tussen haakjes zijn geclusterd op aanbiedersniveau. Gebaseerd op een tweezijdige test geven *, ** en *** respectievelijk de significantie op 0,10, 0,05 en 0,01 niveau weer.

Model (b) en (c) laten zien dat er verschillen bestaan tussen hoofddiagnoses. De geschatte kans ligt 18,8 procentpunt hoger bij ontwikkelingsstoornissen en 33,2 procentpunt bij persoonlijkheidsstoornissen. In de algemene regressievergelijking veronderstellen we dat de kans op een toename in de GAF-score onafhankelijk is van de start GAF-score. In model (e) en (f) kijken we daarom alleen naar behandelingen met dezelfde start GAF-score. De geschatte kans op verbetering is in beide gevallen hetzelfde, en is 26,8 procentpunt hoger bij vrijgevestigde behandelaren.

Discussie

De schattingsresultaten laten zien dat de kans op een toename in de GAF-score substantieel hoger is, namelijk meer dan anderhalf keer hoger, na behandeling door een vrijgevestigde behandelaar dan na behandeling door een ggz-instelling. Er is hierbij gecorrigeerd voor een groot aantal patiëntkenmerken. De resultaten zijn ook robuust voor verschillende specificaties van het lineaire regressiemodel.

Dit suggereert dat een behandeling bij vrijgevestigde behandelaren eerder tot een toename in de GAF-score leidt dan een behandeling bij ggz-instellingen. Er zijn verschillende verklaringen mogelijk voor deze toename.

Een verklaring is een verschil in kwaliteit tussen beide typen behandelaren. De eerder genoemde verschillen in betaling en zelfselectie van behandelaren op basis van pull en push-factoren kunnen de oorzaak vormen voor dit kwaliteitsverschil. Bijvoorbeeld, in ggz instellingen spelen budgetrestricties een belangrijke rol wat gevolgen kan hebben voor de kwaliteit. Daardoor is er mogelijk onvoldoende capaciteit en tijd beschikbaar om patiënten te behandelen, kan er een tekort zijn aan behandelaren of is er een prikkel om (dure)

psychologen en psychiaters minder in te zetten tijdens een behandeling. Dit speelt veel minder een rol bij vrijgevestigde behandelaren, die per behandeling betaald kregen. In vrijgevestigde praktijken wordt relatief meer behandeld door psychiaters en psychologen, terwijl instellingen vaker gebruikmaken van maatschappelijk werkers, verpleegkundigen of andere therapeuten. Daarnaast besteden vrijgevestigde behandelaren meer directe tijd aan de patiënt en minder indirecte tijd aan het coördineren of aansturen van een multidisciplinair team, zoals behandelaren in ggz-instellingen. Een rol kan daarbij spelen dat de vrijgevestigde behandelaar een betere band kan opbouwen met de patiënt, omdat de kans dat deze dezelfde behandelaar blijft zien groter is. Een andere reden voor het kwaliteitsverschil kan zijn een verschil in specialisatie. Behandelaren in ggz-instellingen krijgen meer mensen met zware psychische klachten, met een crisisopname of mensen die meerdere complexe en langdurige behandelingen nodig hebben en zullen zich mogelijk meer in specialiseren in die type klachten. Vrijgevestigde behandelaren krijgen vooral mensen met lichtere psychische klachten en zullen zich vooral in die problematiek specialiseren. Omdat ons onderzoek zich vooral richt op lichtere psychische klachten zou specialisatie een mogelijke verklaring kunnen zijn waarom we in ons onderzoek een grotere kans op verbetering van de GAF-score vinden bij vrijgevestigde behandelaren. Deze reden impliceert dat betere behandeluitkomsten kunnen worden bereikt wanneer behandelaren zich specialiseren in type psychische klachten.

Een tweede mogelijke verklaring is dat vrijgevestigde behandelaren patiënten bij aanvang (beëindiging) van de behandeling lager (hoger) scoren op de GAF dan behandelaren in dienst van ggz-instellingen. Vrijgevestigde behandelaren kunnen bijvoorbeeld andere normen aanhouden bij het invullen van GAF-scores (Greenberg en Rosenbeck, 2005). Ze kunnen optimistischer zijn bij het beëindigen van de behandeling over het psychisch, sociaal, en beroepsmatig functioneren van hun patiënten. Zo kunnen ze aan het einde van een behandeling een hogere GAF-score geven, omdat ze meer tevreden zijn over hun eigen behandeling. Vrijgevestigde behandelaren behandelen de patiënt immers veelal zelf, terwijl er bij een ggz-instelling vaker verschillende behandelaars zijn. Omgekeerd kunnen behandelaren in dienst van een ggz-instelling pessimistischer zijn. Zij zien immers ook regelmatig mensen met zwaardere psychische klachten die na een behandeling terugkeren. Hierbij speelt mogelijk een rol dat bij een instelling de behandelaar die de GAF-score aan het einde van de behandeling invult niet dezelfde behandelaar hoeft te zijn als de behandelaar die de GAF-score invult aan het begin van de behandeling. Hierdoor kunnen de scores mogelijk lager uitvallen dan wanneer dezelfde behandelaar de GAF-score invult aan het begin en einde van een traject. Het is ook mogelijk dat er systematische verschillen in GAF-scores ontstaan tussen beide typen aanbieders, omdat behandelaren geen direct belang hebben bij het nauwkeurig invullen van de GAF-score. Enerzijds heeft het niet hebben van direct belang het voordeel dat strategisch scoren, zoals genoemd door Greenberg en Rosenbeck (2005), niet optreedt, maar anderzijds heeft dit als nadeel dat het risico op onzorgvuldig invullen van GAF-scores wordt vergroot. Onzorgvuldig invullen van GAF-scores kan ook leiden tot systematische verschillen tussen behandelaren.

Een derde mogelijke verklaring is dat, ondanks alle correcties van beschikbare karakteristieken, patiënten die door de vrijgevestigde behandelaren worden behandeld niet volledig vergelijkbaar zijn met de patiënten die door een ggz-instelling worden behandeld. Een mogelijke reden voor dit verschil is dat er complexe patiënten zijn in onze populatie die alleen door een ggz-instelling mogen worden behandeld. Ook kunnen vrijgevestigde behandelaren besluiten om complexe patiënten door te sturen naar een ggz-instelling (Bannink et al., 2020). Selectie door patiënten of behandelaren kan ook een rol spelen. Zo kunnen bepaalde typen patiënten de voorkeur geven aan een individuele behandeling bij een vrijgevestigde behandelaar. De samenstelling van de onderzoekspopulatie en de controlevariabelen houden hier weliswaar tot op zekere hoogte rekening mee, maar het is mogelijk dat dit toch onvoldoende is. De toewijzing van patiënten aan DBC's kan ook verschillen tussen vrijgevestigde behandelaren en ggz-instellingen vanwege een verschil in inschatting in de diagnose. Ook worden patiëntkarakteristieken als opleiding, inkomen, en migratieachtergrond niet direct meegenomen in de analyse. Als mensen zonder migratieachtergrond, met een hogere opleiding en een hoger inkomen een hogere kans hebben om vooruit te gaan, en vaker voor een vrijgevestigde behandelaar kiezen, dan kan dit (een deel van) het verschil verklaren.

Conclusie

Voor de aansturing van de GGZ is het belangrijk om de effectiviteit van behandelingen goed te kunnen beoordelen. In de GGZ is dit een probleem omdat er weinig inzicht is in de effectiviteit van behandelingen. Uit onze analyse volgt dat de kans op verbetering in het sociaal, psychisch en beroepsmatig functioneren van mensen met psychische klachten, gemeten met GAF-scores, hoger is na een behandeling door een vrijgevestigde behandelaar dan door een behandelaar in dienst bij een ggz-instelling.

We vinden een verschil in GAF-score van 27,4 procentpunt waarbij we rekening houden met verschillen in de patiëntpopulatie. Dit verschil is substantieel en suggereert dat de kans op een toename van de GAF-score van een patiënt meer dan anderhalf keer hoger is bij een vrijgevestigde aanbieder dan bij een ggz-instelling. Dit kan een mogelijke aanwijzing zijn dat de kwaliteit van een behandeling, zoals gemeten door de GAF-score, hoger is bij vrijgevestigde behandelaren dan bij behandelaren in een ggz-instelling.

We kunnen echter niet concluderen dat er een causaal verband is in behandeluitkomsten tussen een vrijgevestigde behandelaren en een behandelaar in een ggz-instelling. Ten eerste is de GAF-score een beperkte maat, want deze meet slechts een bepaald aspect van de behandeluitkomst. Ten tweede zijn er alternatieve verklaringen voor het gevonden verschil in GAF-score. Een alternatieve verklaring is dat vrijgevestigde behandelaren de GAF-scores systematisch anders invullen dan behandelaren in dienst bij een ggz-instelling.

Dit artikel moet daarom vooral gezien worden als een uitnodiging tot het doen van verder onderzoek over dit maatschappelijk relevante onderwerp. Meer inzicht in de kwaliteit van een behandeling is te verkrijgen met andere uitkomstmaten zoals ROM ('Routine Outcome

Measuring), PROMS ('Patient Reported Outcome Measures'), de HONOS ('Health of the Nation Outcomes Scale') of met bredere economische uitkomstmaten die samenhangen met gezondheid zoals bijvoorbeeld inkomen en arbeidsparticipatie.

Auteurs

Rudy Douven (e-mail: R.Douven@cpb.nl) is onderzoeker aan de Erasmus Universiteit Rotterdam en wetenschappelijk medewerker bij het Centraal Planbureau; Tobias Hlobil (e-mail: t.m.hlobil@cpb.nl) is wetenschappelijk medewerker bij het Centraal Planbureau en Lisanne Schippers (e-mail: lisanne.schippers@outlook.com) was tijdens het onderzoek master student aan de Erasmus Universiteit Rotterdam en liep stage bij het Centraal Planbureau.

Literatuur

- Angrist, J.D. en J.S. Pischke, 2008, *Mostly Harmless Econometrics*, Princeton University Press.
- Bannink, D., J. de Groot en H. Taverne, 2020, Ex ante beleidsevaluatie Zorgprestatie model ggz en fz. TALMA instituut. <https://www.zorgprestatie.nl/shared/content/uploads/2020/07/Talma-Evaluatie-Zorgprestatie-model-GGZ-en-FZ.pdf>
- Chandra, A., D. Cutler en Z. Song, 2012), Who ordered that? The economics of treatment choices in medical care, in: M.V. Pauly, T.G. McGuire en P.P.Barros (eds), *Handbook of Health Economics*, volume 2, chapter 6: 397-432.
- Douven, R., Remmerswaal en I. Mosca, 2015, Unintended effects of reimbursement schedules in mental health care, *Journal of Health Economics*, vol. 42: 139-150. <https://doi.org/10.1016/j.jhealeco.2015.03.008>
- Douven, R., M. Remmerswaal en T. Vervliet, 2021, Payment schemes and treatment responses in mental, *Health Economics*, online 07 september 2021. <https://doi.org/10.1002/hec.4417>
- Douven, R., M. Remmerswaal en R. Zoutenbier, 2017, Do Altruistic Mental Health Care Providers Have Better Treatment Outcomes?, *Journal of Human Resources*, vol. 54(2): 310-341. <https://doi.org/10.3368/jhr.54.2.0716.8070r1>
- Greenberg, G.A. en R.A. Rosenheck, 2005, Using the GAF as a national mental health outcome measure in the Department of Veterans Affairs, *Services*, vol. 56(4):420-426. <https://doi.org/10.1176/appi.ps.56.4.420>
- Grootenboer, E.M.V., Giltay, E.J., van der Lem, R., van Veen, T., van der Wijje, N.J.A. en F.G. Zitman, 2012, Reliability and validity of the Global Assessment of Functioning Scale in clinical outpatients with depressive disorders. *Journal of Evaluation in Clinical Practice*, 18(2): 502-507. <https://doi.org/10.1111/j.1365-2753.2010.01614.x>
- Hellevik, O., 2007, Linear versus logistic regression when the dependent variable is a dichotomy, *Qual Quant*, vol. 43: 59-74. <https://doi.org/10.1007/s11135-007-9077-3>
- Jong, K. de, 2016, Behandelresultaten van vrijgevestigde psychologen en Psychotherapeuten, *Tijdschrift voor Psychotherapie*, vol. 42(5): 308-317. <https://doi.org/10.1007/s12485-016-0148-6>
- Joldersma, C., 2019, maart, *Psychiaters in ggz-instellingen: Arbeidsmarktonderzoek NVvP-LAD*. <https://www.lad.nl/wp-content/uploads/2019/03/Arbeidsmarktonderzoek-psychiaters-in-ggz-instellingen-21032019.pdf>

- Jongebreur W., M. Schipper, L. Vunderink en M. Zwanepol, 2015, Het in beeld brengen van kwaliteit van leven in de geestelijke gezondheidszorg, Significant rapport [Het in beeld brengen van kwaliteit van leven in de geestelijke gezondheidszorg \(rapport Significant\) | Rapport | Zorginstituut Nederland](#)
- NZa, 2013, 27 juni, Tariefbeschikking twijdelijns curatieve GGZ - TB/CU-5066-02-volgnr.2 - Nederlandse Zorgautoriteit. Nederlandse Zorgautoriteit. https://puc.overheid.nl/nza/doc/PUC_5701_22/1/
- NZa, (2014, 1 oktober, Gespecialiseerde GGZ - NR/CU-552 - Nederlandse Zorgautoriteit. Publicatieplatform UitvoeringsContent Nederlandse Zorgautoriteit. https://puc.overheid.nl/nza/doc/PUC_1834_22/1/
- NZa, 2019, Stand van de zorg 2019. Aan de slag. Nederlandse Zorgautoriteit. https://puc.overheid.nl/nza/doc/PUC_291461_22/1
- Söderberg, P., S. Tungström en B.A. Armelius, 2005, Special section on the GAF: reliability of Global Assessment of Functioning ratings made by clinical psychiatric staff, *Psychiatric Services*, vol. 56(4): 434-438. <https://doi.org/10.1176/appi.ps.56.4.434>
- Söderberg, P. en S. Tungström, 2007, Outcome in psychiatric outpatient services: Reliability, validity and outcome based on routine assessments with the GAF scale. Doctoral dissertation from the Department of Psychology, Umeå University, S-901 87 UMEÅ, Swijden: ISBN 91-7264-240-8; ISBN 978-91-7264-240-9
- Stefanopoulou, E., A.R. Lafuente, A.S. Fonseca, S. Keegan, C. Vishnick en A. Huxley, 2011, Global assessment of psychosocial functioning and predictors of outcome in schizophrenia, *International journal of psychiatry in clinical practice*, vol. 15(1): 62-68. <https://doi.org/10.3109/13651501.2010.519035>
- Urbanoski, K.A., C. Henderson en S. Castel, 2014., Multilevel analysis of the determinants of the global assessment of functioning in an inpatient population, *BMC Psychiatry*, vol. 14(1): 1-6 <https://doi.org/10.1186/1471-244X-14-63>
- Vektis, 2017, november, *Zorgthermometer ggz | Inzicht in de geestelijke gezondheidszorg*. <https://www.vektis.nl/intelligence/publicaties/zorgthermometer-inzicht-in-de-ggz>
- Verhaak, P.F.M., H. Kamsma en A. van der Niet, 2013., Mental Health Treatment Provided by Primary Care Psychologists in the Netherlands, *Psychiatric Services*, vol. 64(1): 94-97. <https://doi.org/10.1176/appi.ps.201100541>
- Warmerdam, E., P. ten Have, J. Dekker en E. de Beurs, 2016, Klachtenniveau bij aanvang behandeling is belangrijkste voorspeller van behandeluitkomst bij angst en depressie, *GZ - Psychologie*, vol. 8(4): 10-15. <https://doi.org/10.1007/s41480-016-0026-4>
- Warr, P. en I. Inceoglu, 2018, Work Orientations, Well-Being and Job Content of Self-Employed and Employed Professionals, *Work, Employment and Society*, vol. 32(2): 292-311. <https://doi.org/10.1177/0950017017717684>
- ZiN., 2020, December, Landelijk kwaliteitsstatuut GGZ. <https://www.zorginzicht.nl/kwaliteitsinstrumenten/ggz-landelijk-kwaliteitsstatuut>
- Zoutenbier, R. en R.. Douven, 2016, Relatie behandelduur en GAF-scores in de GGZ, CPB Notitie 23 februari 2016. <https://www.cpb.nl/publicatie/relatie-behandelduur-en-gaf-scores-in-de-ggz>