

# Gedragsreacties van ggz-aanbieders op financiële prikkels

In 2008 is prestatiebekostiging ingevoerd voor vrijgevestigde zorgaanbieders in de curatieve geestelijke gezondheidszorg. Onderzoek op basis van administratieve data laat zien dat sommige vrijgevestigde aanbieders sterker reageren op de prestatiebekostiging dan andere. Aanbieders die sterker reageren behandelen gemiddeld genomen langer maar rapporteren geen betere behandeluitkomsten.

## RUDY DOUVEN

Wetenschappelijk medewerker bij het Centraal Planbureau (CPB) en onderzoeker aan de Erasmus Universiteit Rotterdam

## MINKE REMMERSWAAL

Wetenschappelijk medewerker bij het CPB

## ROBIN ZOUTENBIER

Wetenschappelijk medewerker bij het CPB

Verschillen in motivaties van behandelaren in de zorg kunnen bijdragen aan het verklaren van variatie in uitkomsten in de zorg. Dit geldt met name in de geestelijke gezondheidszorg (ggz), waar meer onzekerheid bestaat over diagnose en behandelwijze (Frank en McGuire, 2000). Het is voor behandelaren lastig om aan te tonen welke behandeling het meest effectief zal zijn en hoeveel behandeling er nodig is.

In de literatuur worden twee belangrijke bronnen van motivatie van behandelaren aangewezen: de wil om bij te dragen aan de gezondheid van een patiënt en de financiële beloning voor het leveren van zorg. Campbell *et al.* (2012) laten zien dat behandelaren in de zorg de eerste van de twee als belangrijkste motivatie rapporteren. Echter, wanneer economische belangen een rol spelen, zijn behandelaren ook geneigd hun behandelwijze daarop aan te passen.

Veel onderzoek laat zien dat zorgaanbieders reageren op financiële prikkels (Chandra *et al.*, 2012). Ook in Nederland komt er meer empirische evidentie beschikbaar. Vrijgevestigde specialisten in ziekenhuizen blijken vaker te behandelen dan specialisten in loondienst (Douven *et al.*, 2015a) en huisartsen pasten hun behandelwijze aan op de invoering van nieuwe financieringssystemen in 2006 (Van Dijk *et al.*, 2013). Eerder onderzoek laat zien dat vrijgeves-

tigde aanbieders die gefinancierd werden via prestatiebekostiging langer behandelen dan aanbieders in loondienst van de voormalig gebudgetteerde zorginstellingen (Douven *et al.*, 2015b). Dat onderzoek richtte zich echter enkel op gemiddelde gedragsreacties van aanbieders en niet op de behandeluitkomsten. In dit onderzoek laten we zien dat er grote verschillen kunnen bestaan in de extrinsieke motivaties van aanbieders, en dat deze verschillen van invloed zijn op behandeluitkomsten en zorgkosten. Deze inzichten bieden een leidraad bij de verdere ontwikkeling van prestatiebekostiging in de zorg.

## PRESTATIEBEKOSTIGING IN DE GGZ

In 2008 zijn er delen van de ggz overgeheveld vanuit de Algemene wet bijzondere ziektekosten (AWBZ) naar de Zorgverzekeringswet (Zvw). Dit is de curatieve ggz geworden waarbij er een onderscheid is gemaakt tussen de gespecialiseerde ggz en de generalistische basis-ggz. De gespecialiseerde ggz behelst specialistische curatieve zorg voor patiënten met een psychiatrische aandoening, zoals depressie of schizofrenie. De gespecialiseerde ggz omvat voornamelijk ambulante zorg en behandeling met tijdelijk verblijf in een ggz-instelling. Patiënten met een minder complexe zorgvraag kunnen terecht bij de generalistische basis-ggz. Patiënten kunnen van beide typen ggz gebruikmaken na een verwijzing.

Om de doelmatigheid in de curatieve ggz te verbeteren werd in 2008 prestatiebekostiging ingevoerd. De prestatiebekostiging gold in eerste instantie alleen voor de vrijgevestigde aanbieders, zoals zelfstandige psychiaters en psychologen, en instellingen die na 1 januari 2008 de markt betraden. Deze aanbieders zijn goed voor ongeveer tien procent van de totale omzet in de curatieve ggz (NZa, 2012). De overige, veelal grotere, zorginstellingen kregen een budget toegewezen en vielen niet onder de prestatiebekostiging. Sinds 2013 geldt de prestatiebekostiging voor alle ggz-aanbieders.

Bij prestatiebekostiging declareren aanbieders op basis van een diagnose behandelcombinatie (DBC). Een DBC legt het gehele zorgtraject dat een patiënt doorloopt vast, vanaf het eerste (telefonisch) contact met de behandelaar tot de laatste administratieve handeling. Zorgaanbieders krijgen een vast tarief per DBC (tenzij hierover contractueel andere afspraken zijn gemaakt met de verzekeraar). Dit stimuleert in principe doelmatige zorg, omdat aanbieders baat hebben bij het drukken van de kosten van een behandeling.

## DE PRIKKELS VAN PRESTATIEBEKOSTIGING

In de prestatiebekostiging in de ggz hangt de beloning voor een DBC af van de gestelde diagnose en de behandelduur in minuten; des te langer de behandeling van een patiënt des te hoger het beloningstarief. Het tarief voor een DBC volgt de behandelduur via een trapsgewijze functie; op de tijdsgrenzen van 250, 800, 1800, 3000 en 6000 minuten stijgt het tarief voor een DBC. Figuur 1 toont als voorbeeld de tariefstructuur van de DBC 'depressie'.

De prestatiebekostiging geeft een sterke financiële prikkel. Als een zorgaanbieder een patiënt bijvoorbeeld 1801 minuten heeft behandeld, dan krijgt de aanbieder geen extra financiële vergoeding voor langer doorbehandelen tot de volgende tijdsgrens van 3000 minuten. De aanbieder zal dus na 1801 minuten alleen doorgaan met het behandelen van de patiënt wanneer de baten (hier: het gezonder worden van de patiënt of de reputatie-effecten die daar van uitgaan) opwegen tegen de kosten die de aanbieder moet maken (de extra tijd die naar de patiënt toegaat). Dit is het doelmatigheidseffect van de prestatiebekostiging en het voorkomt dat een aanbieder onnodig lang blijft doorbehandelen.

De trapsgewijze functie kent ook een onbedoeld effect. Net voor de tijdsgrenzen is er een prikkel voor aanbieders om langer door te behandelen dan nodig. Een aanbieder die na 1700 minuten behandelen stopt ontvangt een tarief van 2050 euro, wanneer de aanbieder de behandeling verlengt met 100 minuten stijgt het tarief tot 3703 euro. Langer doorbehandelen levert dus een forse verhoging in het tarief op.

## ONDERZOEKSPOPZET

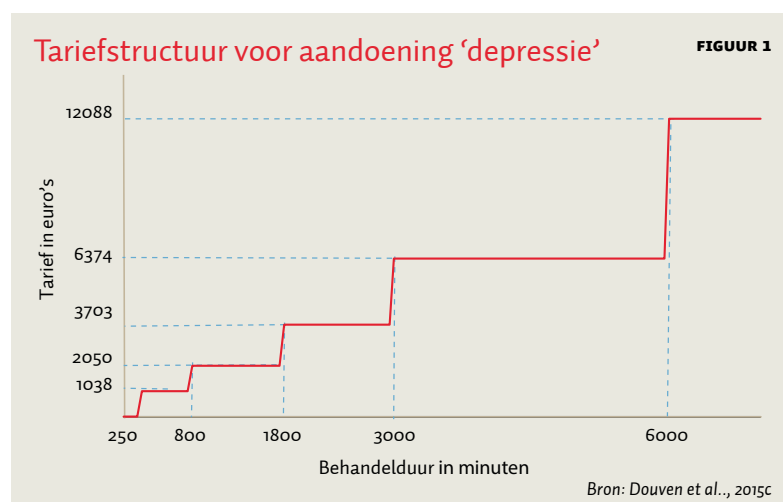
We hebben voor iedere afzonderlijke vrijgevestigde aanbieder onderzocht in welke mate ze strategisch reageren op de trapsgewijze tarieven en of er verschillen zijn tussen aanbieders (Douven *et al.*, 2015c). De bestudeerde gegevens komen uit het DBC Informatiesysteem (DIS). Deze database bevat informatie over de behandelingen en aanbieders in de curatieve ggz over de periode 2008 tot en met 2010. Onze sample bevat 1440 vrijgevestigde aanbieders die in deze periode samen 188.000 DBC's hebben uitgevoerd. Verder kijken we alleen naar DBC's van maximaal 4000 minuten (langere DBC's zijn zeldzaam) en alleen naar reguliere DBC's, dus bijvoorbeeld geen korte DBC's (die meestal korter zijn dan 400 minuten en een andere tariefstructuur kennen).

De aanbieders zijn opgedeeld in aanbieders die sterk reageren op de financiële prikkels in de prestatiebekostiging en aanbieders dat niet doen. Figuur 2 geeft voorbeelden van de verdeling van behandelduur voor drie afzonderlijke aanbieders. Figuur 2a toont een aanbieder die niet reageert

op de financiële prikkel. De verdeling van behandelduur toont geen strategisch gedrag rondom de drie tariefgrenzen van 800, 1800 en 3000 minuten. We classificeren deze aanbieder daarom als niet of minder extrinsiek gemotiveerd. De aanbieders in figuur 2b en 2c vertonen – in toenemende mate – wel strategisch gedrag rondom de tariefgrenzen. Beide figuren tonen meer afgesloten DBC's na de tijdsgrenzen dan ervoor. Het strategisch gedrag is het grootst bij de aanbieder in figuur 2c, een groot deel van de DBC's wordt net na de tariefgrens gesloten. De aanbieders die vaker strategisch gedrag vertonen, zoals in figuur 2b en 2c zijn geclassificeerd als meer extrinsiek gemotiveerd.

Alle 1440 vrijgevestigde aanbieders zijn op deze wijze geclassificeerd, als minder extrinsiek gemotiveerd of als meer extrinsiek gemotiveerd. Hiervoor hebben we voor elke aanbieder geteld hoeveel DBC's er gesloten zijn voor een tariefgrens en hoeveel na een tariefgrens. Aanbieders met relatief meer afgesloten DBC's na de tariefgrens en relatief minder voor de tariefgrens zijn geclassificeerd als meer extrinsiek gemotiveerd. Aanbieders met relatief meer DBC's afgesloten voor de tariefgrens en minder na de tariefgrens zijn minder extrinsiek gemotiveerd. Er blijkt grote variatie tussen aanbieders te zijn. Ongeveer 25 procent van de aanbieders lijken sterk op figuur 2a en zijn minder extrinsiek gemotiveerd. Van de overige 75 procent van de aanbieders is het merendeel meer extrinsiek gemotiveerd (figuren 2b en 2c).

Wellicht wordt de variatie tussen aanbieders veroorzaakt door verschillen in patiëntenpopulatie (sommige aanbieders kunnen bijvoorbeeld meer patiënten met zwaardere diagnoses hebben) of verschillen in eigenschappen van aanbieders (er kunnen verschillen zijn tussen behandelmethodes van psychiaters of psychologen). Dit bleek niet het geval te zijn. Uit onze data blijken er geen grote verschillen zijn in observeerbare patiënt- en aanbiedereigenschappen tussen de groep extrinsiek gemotiveerde aanbieders en de groep minder extrinsiek gemotiveerde aanbieders. Daarnaast zien we ook geen concentratie van extrinsieke aanbieders op provincieniveau. De empirische analyses corrigeren zo goed mogelijk voor verschillen in patiënt- en aanbiedereigenschappen (zoals gemiddelde leeftijd en geslacht) tussen extrinsiek en minder extrinsiek gemotiveerde aanbieders. Het kan zijn dat onze resultaten worden be-



invloed door patiëteigenschappen die niet te observeren zijn, maar de richting van dat effect is a priori niet duidelijk (Douven *et al.*, 2015c).

### LANGERE BEHANDELINGEN

Extrinsiek gemotiveerde aanbieders behandelen langer. Of de meer extrinsiek gemotiveerde aanbieders gemiddeld genomen langer of korter behandelen is a priori niet duidelijk. Enerzijds kunnen ze eerder stoppen met behandelen dan de minder extrinsiek gemotiveerde aanbieders. Bijvoorbeeld als extrinsiek gemotiveerde aanbieders stoppen zodra ze een tijdsgrens passeren, terwijl minder extrinsiek gemotiveerde aanbieders doorbehandelen. Anderzijds kunnen meer extrinsiek gemotiveerde aanbieders ook langer doorbehandelen, bijvoorbeeld omdat ze de volgende tijdsgrens met een hoger tarief willen bereiken.

Uit de analyse blijkt dat de meer extrinsiek gemotiveerde aanbieders gemiddeld 172 minuten langer behandelen dan minder extrinsiek gemotiveerde aanbieders. Dus waar een minder extrinsiek gemotiveerde aanbieder kiest te stoppen met behandelen, blijkt een extrinsiek gemotiveerde aanbieder gemiddeld genomen bijna drie uur langer door te behandelen tot een volgend tarief is bereikt. De empirische analyse laat ook zien dat naarmate aanbieders sterker extrinsiek gemotiveerd zijn ze eerder, of vaker, besluiten om langer door te behandelen.

Het verschil in behandelduur heeft ook een effect op de zorguitgaven. Een gemiddelde DBC van de meer extrinsiek gemotiveerde aanbieders is gemiddeld ongeveer twintig procent (ongeveer 400 euro) duurder dan een gemiddelde DBC van minder extrinsiek gemotiveerde aanbieders. Het verschil in de gemiddelde prijs per uur tussen meer extrinsiek en minder extrinsiek gemotiveerde aanbieders is met vijf procent (ongeveer 5 euro per uur) minder groot. Er kunnen meerdere redenen zijn voor dit kleinere verschil. Allereerst zijn de tarieven zo vastgesteld dat behandelingen korter dan 800 minuten gemiddeld meer opleveren per uur. Omdat minder extrinsiek gemotiveerde aanbieders vaker korter behandelen, stijgt hun gemiddelde uurtarief. Ten tweede is het mogelijk dat meer extrinsiek gemotiveerde aanbieders niet alleen doorbehandelen naar een volgende

tariefgrens, maar ook naar een daaropvolgende tariefgrens. In dat geval nemen de behandelduur en het tarief van een DBC toe terwijl het gemiddeld uurtarief niet stijgt.

### GEEN BETERE UITKOMSTEN

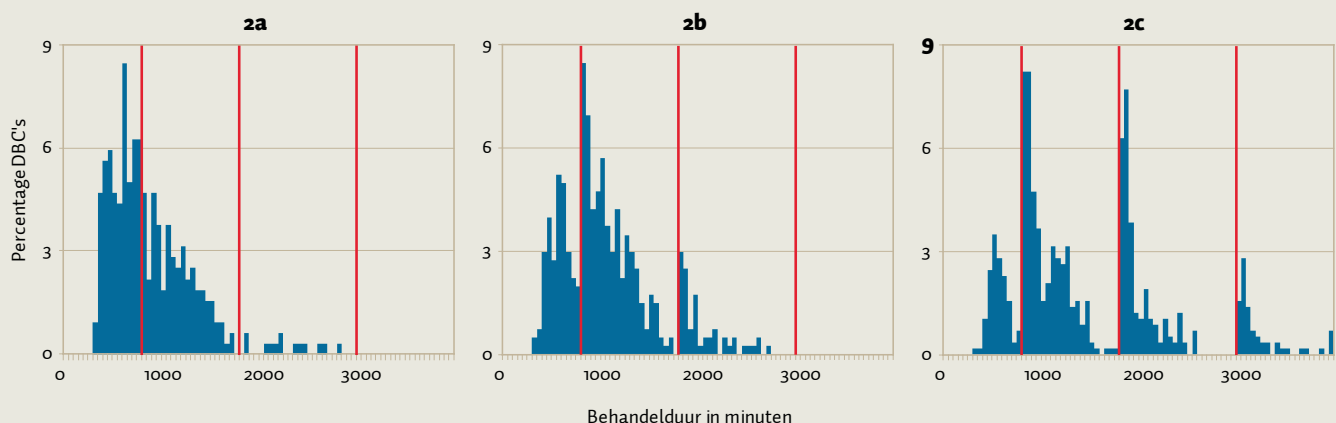
Het is moeilijk om te bepalen wat wenselijk is vanuit het algemene welvaartsbelang: langer of korter behandelen. Langer behandelen is weliswaar duurder maar kan de welvaart verhogen omdat de gezondheid van de patiënt toeneemt. We vergelijken daarom de gezondheidsuitkomsten van de behandelingen van extrinsiek en minder extrinsiek gemotiveerde aanbieders. Het meten van gezondheidsuitkomsten is een van de lastigste aspecten in de economie van de gezondheidszorg, zeker in de geestelijke gezondheidszorg.

Om een beeld te krijgen van de uitkomst van een behandeling vergelijken we de zogenoemde Global Assessment of Functioning (GAF). De GAF-score is een getal tussen 1 en 100 die aangeeft hoe een patiënt functioneert in het dagelijks leven. Hoe lager de GAF-score, hoe slechter de patiënt in psychisch, sociaal of beroepsmatig opzicht functioneert. Een GAF-score van 1 betekent bijvoorbeeld dat iemand dagelijks last heeft van suïcidegedachten en een GAF-score van 100 betekent dat een patiënt perfect functioneert in het dagelijkse leven. De behandelaar scoort de patiënt zowel aan het begin van de behandeling als aan het eind van de behandeling. Het verschil in GAF-score tussen het begin en eind van de behandeling toont de ontwikkeling aan gedurende de behandeling en geeft daarmee informatie over de uitkomst van een behandeling. Alle zorgaanbieders in de ggz zijn bekend met het afnemen van de GAF-score bij de patiënt en rapporteren deze al jaren routinematig.

Er is ook kritiek op het gebruik van de GAF-score. In eerste instantie meet de GAF-score alleen hoe iemand functioneert in het dagelijkse leven. Dit is slechts één dimensie van de gezondheidsstatus van een patiënt. Voor onderzoek naar kwaliteit van zorg is het beter om ook naar andere kwaliteitsindicatoren te kijken, zoals bijvoorbeeld ROM scores. Deze zijn echter nog niet op grote schaal beschikbaar. Daarnaast is ook onderzoek gedaan naar de betrouwbaarheid van GAF-scores (Jones *et al.*, 1995, Soderberg *et al.*, 2014). Deze studies laten zien dat de GAF-

Verdeling van behandelduur

FIGUUR 2



Bron: Douven *et al.*, 2015c

scores vooral betrouwbaar zijn wanneer grote groepen patiënten worden vergeleken, zoals in ons onderzoek.

Schattingen waarbij we corrigeren voor patiënt- en aanbiedereigenschappen, laten zien dat minder extrinsiek gemotiveerde aanbieders gemiddeld genomen 1 GAF-punt (op een schaal van 100) meer verbetering van een patiënt rapporteren dan extrinsiek gemotiveerde aanbieders. Dit verschil is significant. Over de gehele populatie is de gemiddelde toename in GAF-score ongeveer 10 punten, een verschil van 1 punt komt dus overeen met een tien procent hogere toename. Daarnaast is ook gecontroleerd of de patiënten van beide type aanbieders niet verschillen in GAF-score bij de start van een behandeling. De gemiddelde GAF-score aan het begin van een behandeling is ongeveer gelijk bij extrinsiek en minder extrinsiek gemotiveerde aanbieders. Dit geeft aan dat de patiënten bij binnenkomst niet verschillen in hun mate van functioneren en dat zorguitkomsten van beide type aanbieders niet vertekend worden door verschillen in de ernst van de psychiatrische aandoening van patiënten.

Deze uitkomsten tonen dat de langere behandelduur bij extrinsiek gemotiveerde aanbieders niet tot betere zorguitkomsten leidt maar eerder tot lagere uitkomsten dan die van de minder extrinsiek gemotiveerde aanbieders. Daar komt bij dat de kosten van een DBC ook hoger liggen dan bij de minder extrinsiek gemotiveerde aanbieders. Op meerdere manieren is gekeken naar de gevoeligheid van de resultaten. Bijvoorbeeld door uitkomsten opnieuw te berekenen voor aanbieders met veel DBC's, aanbieders met weinig DBC's of door aanbieders die in zeer beperkte mate extrinsiek gemotiveerd zijn uit de analyse te laten. Een andere belangrijke uitkomst van deze analyses is dat naarmate aanbieders meer extrinsiek gemotiveerd zijn, de behandelduur toeneemt en de verbetering in de GAF-score afneemt.

## CONCLUSIE EN CONSEQUENTIES VOOR BELEID

De analyse van de prestatiebekostiging in de ggz laat zien dat zorgaanbieders verschillend reageren op financiële prikkels. Sommige aanbieders laten zich leiden door financiële prikkels. Ze krijgen meer betaald, maar rapporteren niet per se betere uitkomsten. Ongeveer een kwart van de aanbieders reageert niet of nauwelijks op de sterke financiële prikkels in de bekostiging. Hiervoor zijn verschillende verklaringen mogelijk. Ten eerste zijn sommige aanbieders mogelijk niet bewust van de financiële prikkel of konden ze de precieze werking van de financiële prikkel niet overzien. Een alternatieve verklaring is dat de minder extrinsiek gemotiveerde aanbieders ook intrinsiek gemotiveerd zijn en vooral nut halen uit het beter maken van de patiënt, onafhankelijk van de financiële beloning. In dat laatste geval zullen beleidsmatige aanpassingen in de financiering weinig effect sorteren voor minder extrinsieke zorgaanbieders.

Het is onduidelijk of prestatiebekostiging beter of slechter is dan budgettering of volledig variabele beloning. De trapsgewijze tariefstructuur maakt het eenvoudig om de strategische reacties helder in kaart te brengen, waar dit bij andere vormen van bekostiging vaak moeilijker of zelfs onmogelijk is. Bovendien ontvangen verzekeraars sinds 2014 meer informatie over de gedeclareerde behandelduur. Hierdoor wordt het voor verzekeraars mogelijk om strategische

reacties door aanbieders in kaart te brengen en aanbieders hierop aan te spreken. Verder kregen zorgverzekeraars na 2012 meer onderhandelingsruimte en werd er scherper gecontracteerd. Er worden bijvoorbeeld grotere afslagen op het maximum DBC-tarief gerapporteerd en vaker afspraken met aanbieders gemaakt over omzet- en prijsplafonds. Bij dit laatste aspect speelt ook een rol dat sinds 2013 een macrobeheersinstrument van kracht is in de ggz (NZa, 2015).

In de ggz wordt steeds meer informatie verzameld over de uitkomsten van een behandeling. Voor de toekomst is het van belang dat financieringssystemen meer belonen op basis van uitkomsten dan input-factoren, zoals behandelduur. Aanbieders kunnen dan meer beloond worden op basis van productiviteit. Bij gelijkblijvend gedrag zouden volgens onze studie de minder extrinsiek gemotiveerde aanbieders dan juist meer betaald krijgen en de meer extrinsiek gemotiveerde aanbieders juist minder. Zorgverzekeraars kunnen die informatie ook gebruiken om de verschillen in kwaliteit van zorg in kaart te brengen en beter in te kopen.

De NZa is inmiddels bezig met de ontwikkeling van een nieuw financieringssysteem voor de ggz dat is gebaseerd op het Engelse systeem waarbij er ook rekening wordt gehouden met de uitkomst van een behandeling (NZa, 2015). Aanbieders hoeven dan niet meer het precieze aantal minuten te registreren en dienen vooraf aan te geven welke soort behandeling een patiënt nodig heeft. De brede invoering van dit nieuwe systeem staat gepland voor 2019.

## LITERATUUR

- Campbell, E.G. et al. (2007) Professionalism in medicine: results of a national survey of physicians. *Annals of Internal Medicine*, 147(11), 795–802.
- 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.B. Barros (red.), *Handbook of Health Economics*, volume 2. Amsterdam: Elsevier, 397–432.
- Dijk, C.E. van, B. van den Berg, R.A. Verheij et al. (2013) Moral hazard and supplier-induced demand: empirical evidence in general practice. *Health Economics*, 22(3), 340–352.
- Douven, R., R. Mocking en I. Mosca (2015a) The effect of physician remuneration on regional variation in hospital treatments. *International Journal of Health Economics and Management*, 15(2), 215–240.
- Douven, R., M. Remmerswaal en I. Mosca (2015b) Unintended effects of reimbursement schedules in mental health care. *Journal of Health Economics*, 42, 139–150.
- Douven, R., M. Remmerswaal en R. Zoutenbier (2015c) Do extrinsically motivated mental health providers have better treatment outcomes. *CPB Discussion Paper*, 319. Den Haag.
- Jones, S.H., G. Thornicroft, M. Coffey en G. Dunn (1995) A brief mental health outcome scale-reliability and validity of the global assessment of functioning (GAF). *British Journal of Psychiatry*, 166(5), 654–659.
- NZa (2012) *Marktscan geestelijke gezondheidszorg. Weergave van de markt 2008-2011*. Utrecht: Nederlandse Zorgautoriteit.
- NZa (2015) *Voorrang aan nieuwe productstructuur in ggz en forensische zorg*. Utrecht: Nederlandse Zorgautoriteit.
- Soderberg, P., S. Tungstrom en B.A. Armelium (2014) Special section on the GAF: reliability of global assessment of functioning ratings made by clinical psychiatric staff. *Psychiatric Services*.