

## Benchmarken op basis van ROM-data

Nienke van Sambeek

Al maanden lijkt het bijna nergens anders meer over te gaan in de ggz: Routine Outcome Monitoring (ROM). Zowel via sociale media als vakbladen wordt er hevig gediscussieerd over de manier waarop ROM gebruikt wordt. De pennenstrijd barstte begin dit jaar los na publicatie van een rapport van de Algemene Rekenkamer over bekostiging van de curatieve geestelijke gezondheidszorg (ggz) <sup>1</sup>. Daarin werd onder andere geconcludeerd dat ROM-data nog van onvoldoende kwaliteit zijn om de beoogde uitkomstbekostiging op te baseren. Een aantal zorgprofessionals startte naar aanleiding van dit rapport de petitie StopROM <sup>2</sup>, waarop RedROM <sup>3</sup> als reactie kwam.

Over het gebruik van ROM voor de individuele behandeling is relatief weinig discussie. Zowel voor- als tegenstanders van ROM lijken het erover eens dat ROM als klinisch instrument zinvol kan zijn. Zeker wanneer dit flexibel en op maat kan worden ingezet <sup>4</sup>. De discussie over ROM spitst zich toe op het onderwerp van benchmarking: het op groepsniveau vergelijken van uitkomsten met een referentiewaarde, meestal een gemiddelde. Benchmarken gebeurt al in diverse sectoren, bijvoorbeeld om operatiecijfers van ziekenhuizen te vergelijken, of onderwijsuitkomsten bij scholen. Het primaire doel van benchmarken is om zorgaanbieders spiegelinformatie te bieden die kan leiden tot reflectie en kwaliteitsverbetering <sup>5</sup>. Benchmarken is dus bedoeld als leerinstrument. Zorgaanbieders

met minder goede uitkomsten kunnen zichzelf verbeteren door in gesprek te gaan met zorgaanbieders met betere uitkomsten en zo te leren van elkaar. Daarmee kan de zogenaamde 'praktijkvariatie' worden verkleind. In de ggz bestaan momenteel aanzienlijke verschillen in de uitkomsten van ggz aanbieders <sup>6</sup>. Het is belangrijk om die verschillen te verminderen, zodat cliënten overal de best mogelijke zorg kunnen krijgen.

Critici stellen echter dat ROM-data ongeschikt zijn voor het grootschalig vergelijken van uitkomsten van psychologische behandelingen. Zij pleiten daarom voor het stoppen met benchmarken op basis van ROM-data en de landelijke ingevoerde dataverzameling en -verwerking door Stichting Benchmark GGZ (SBG) <sup>2</sup>.



### Wat zeggen de voorstanders?

- Benchmarken met ROM-data is een goed signaleringsmiddel als startpunt om de kwaliteit van psychologische behandelingen van therapeuten of teams te vergelijken en te kunnen verbeteren.
- Er zijn goede methodes om te controleren op eventuele vervuiling van data.
- Door benchmarkinformatie openbaar te maken krijgen cliënten en financiers meer inzicht in de kwaliteit van zorg. Daar hebben ze recht op.



### Wat zeggen de tegenstanders?

- ROM-data zijn ongeschikt om de kwaliteit van psychologische behandelingen te vergelijken.
- Benchmarken geeft onjuiste, troebele uitkomsten, het maakt de zorg niet transparanter of beter.
- Benchmarken op basis van ROM-data dient vooral het belang van zorgverzekeraars, de cliënt heeft er niks aan.



De discussie over ROM is vooral een discussie geworden die fundamentele verschillen in visies blootlegt. Hoe kijken professionals aan tegen hun vak? Hoe denken ze over professionele autonomie versus toezicht en controle? In hoeverre geloven zij dat de kwaliteit van hun vak meetbaar is? Bij de VGCT zien we het meten van behandeluitkomsten als een belangrijk ingrediënt van doelgericht behandelen en denken we dat het succes van een behandeling zich onder andere vertaalt in een afname van klachten. We zien ook goede mogelijkheden om de opgetelde uitkomsten van individuele behandelingen te

gebruiken voor onderlinge vergelijking en kwaliteitsverbetering. Voor benchmarken worden betrouwbare meetinstrumenten en methodes ingezet die ook in wetenschappelijk onderzoek gebruikt worden en waar cgt-behandelingen hun ‘evidence-based’ karakter aan ontleen. Tegelijkertijd zien we ook dat er beperkingen kleven aan benchmarken op basis van ROM-data. We gaan hieronder in op die verschillende bezwaren en sluiten af met een verkenning van oplossingsrichtingen.

► [Lees hier de korte versie](#)

## Issues met benchmarking

### Methodologie

De initiatiefnemers van stopROM spreken zich kritisch uit over de wetenschappelijke validiteit van ROM als benchmarkinstrument. Een van hun voornaamste, methodologische bezwaren gaat over casemix; het gegeven dat verschillende zorgaanbieders verschillende groepen mensen zien. Denk aan variatie in de demografische gegevens en klinische problematiek van cliënten. Het kan daardoor zijn dat een verschil in uitkomsten tussen zorgaanbieders niet toe te schrijven is aan een verschil in kwaliteit, maar aan het verschil in de behandelde cliëntenpopulatie <sup>7</sup>. De angst van veel professionals is dat instelling X, die voornamelijk mensen met moeilijk behandelbare problematiek in een achterstandsbuurt behandelt, ‘oneerlijk’ wordt vergeleken met instelling Y, die voornamelijk hoogopgeleide mensen met enkelvoudige problematiek behandelt.

In wetenschappelijk onderzoek wordt *by design* voor variatie in de cliëntenpopulatie gecontroleerd; mensen worden random toegewezen aan verschillende condities. Bij het vergelijken van groepen op basis van ROM-data die in de klinische praktijk worden verzameld, is deze mate van controle niet mogelijk <sup>8</sup>. Om te voorkomen dat appels met peren worden vergeleken, wordt

een statistische oplossing gebruikt: de case-mixcorrectie. Dit houdt in dat er een homogene subgroep geselecteerd wordt die op relevante kenmerken identiek is <sup>9</sup>. Nu is het alleen zo dat we niet van alle doelgroepen weten welke kenmerken precies relevant zijn <sup>7</sup> en de kenmerken waarvan we dat wel weten worden niet altijd geregistreerd of meegewogen. Kim de Jong is bestuurslid bij de VGCT en als onderzoeker aan de universiteit van Leiden gespecialiseerd in ROM. Zij licht toe: “Op dit moment zitten er nog bepaalde variabelen niet in de casemixcorrectie, waarvan we uit wetenschappelijk onderzoek wel weten dat ze regelmatig voorspellend zijn, zoals chroniciteit. Meer, maar vooral dus betere informatie, is belangrijk om het casemix probleem te verminderen.” Voor zinnige vergelijkingen is het bovendien belangrijk om alleen om subgroepniveau te vergelijken. Bijvoorbeeld door de uitkomsten bij de groep ‘angststoornissen’ met elkaar te vergelijken. Hele instellingen of praktijken met elkaar vergelijken heeft weinig waarde <sup>3</sup>.



Hoe kijkt de wetenschappelijke raad van SBG eigenlijk tegen de methodologische problemen van ROM aan? Gert Westert is hoogleraar kwaliteit van zorg aan de Radboud Universiteit, en lid van die raad. Hij stelt: “Als je goede methodologische eisen stelt en daaraan voldoet, kun je met rommen onderling vergelijken en leren om te verbeteren. SBG voert uitgebreid tests uit op validiteit en kijkt bijvoorbeeld ook of de aangeleverde data representatief zijn voor de instelling. Dus er wordt wel heel veel statistiek op bedreven en dat vind ik zelf wel overtuigend.”

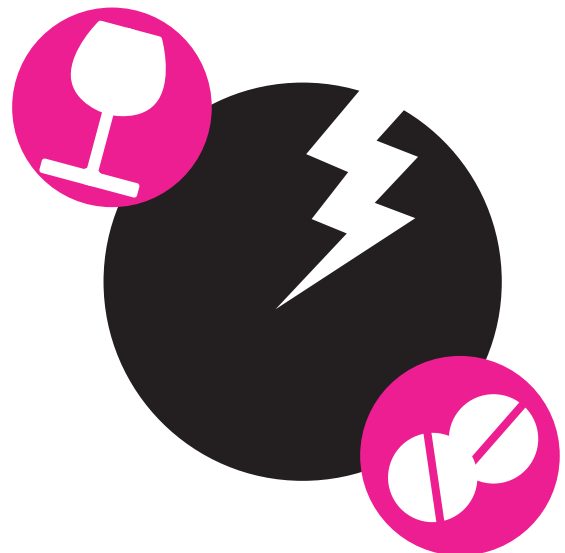
### ***Casemix is vooral een probleem bij EPA-problematiek en verslaving.***

Gevraagd naar het casemixprobleem geeft Westert aan: “Casemix is vooral een probleem bij EPA-problematiek en verslaving, waar mensen chronisch en langdurig in zorg zijn. Dan is het best lastig om goed te voorspellen wat de uitkomsten zouden kunnen zijn van een behandeling en wat die uitkomst eigenlijk moet zijn. Maar als je het hebt over depressie of angststoornissen, dat zijn toch ziektebeelden die beter begrepen worden. Als je bij die stoornissen nog verder corrigeert voor een aantal variabelen, dan kun je heel goed voor casemix corrigeren. Feitelijk doet de SBG dat ook. En doordat de aantallen mensen die voor *common disorders* behandeld worden zo groot zijn, kun je tot goede vergelijkingen komen.”

De validiteit van benchmarking kan dus toe- of afnemen, afhankelijk van de doelgroep waarmee je werkt en de grootte en homogeniteit van die groep. De setting waarin je werkt maakt daarvoor ook uit. ROM-data van vrijgevestigde zorgverleners kunnen niet op dezelfde manier worden gebruikt als data van instellingen. Hans Kamsma, vicevoorzitter van de LVVP legt uit:

“Benchmarken om zelf van te leren, bijvoorbeeld in intervisie, kan ook met data van vrijgevestigden en is zeker zinvol, maar om op grotere schaal verantwoord te kunnen vergelijken heb je data nodig met voldoende zeggingskracht. Daarvoor moet de populatie groot genoeg zijn. Als je dan binnen een kleine praktijk gaat kijken, en ook nog eens gaat corrigeren voor casemix-variabelen, kom je nooit aan aantallen die groot genoeg zijn om echt zeggingskracht te hebben. Dan zijn de individuele afwijkingen echt veel te groot.” Dat is ook de reden dat vrijgevestigde zorgaanbieders alleen op groepsniveau zullen worden vergeleken.

Naast deze puur statistische kwesties van casemix en zeggingskracht plaatsen de initiatiefnemers van stopROM vraagtekens bij de onafhankelijkheid van het onderzoek door SBG, omdat het ‘een door zorgverzekeraars betaalde stichting is en diezelfde zorgverzekeraars ook goed vertegenwoordigd zijn in het bestuur’<sup>2</sup>. Zij vinden dat de database van SBG voor onafhankelijk onderzoek beschikbaar zou moeten zijn. Directeur van SBG Maarten Erenstein erkent desgevraagd dat onafhankelijk onderzoek nodig is voor een onpartijdig oordeel over de kwaliteit van uitkomstinformatie van SBG. Hier worden door de bestuurspartijen van SBG ook stappen in gezet, zo geeft hij aan.



## Benchmarken: met welk doel?

We hebben dus te maken met een grote database aan benchmarkinformatie, die enerzijds een rijke bron van informatie vormt, maar anderzijds nog in ontwikkeling is en methodologische beperkingen kent. Doorontwikkeling en onafhankelijke toetsing moet daar verbetering in brengen. De vraag is nu welk gewicht we aan de genoemde beperkingen moeten toekennen. De experts geven aan dat dit afhangt van het doel waarvoor je de data gebruikt.

Zoals je in deze factsheet ► [link factsheet ROM: waar gaat het heen](#) kunt zien, worden de ROM-benchmarkgegevens momenteel voor verschillende doeleinden tegelijk gebruikt. Voor het leren levert SBG zorgaanbieders benchmarkinformatie op basis van hun ROM-data. Verzekeraars kunnen globaal meekijken in deze rapportages. Verder worden benchmarkgegevens ook voor verantwoording gebruikt. Op instellingsniveau zijn benchmarkgegevens openbare informatie, als onderdeel van de verplichte prestatie-indicatoren die ggz-instellingen aanleveren aan het ministerie van Volksgezondheid, Welzijn en Sport (VWS). Benchmarkgegevens worden op instellingsniveau door het Zorginstituut Nederland (ZiN) openbaar gemaakt<sup>11</sup>, zodat ze ook beschikbaar zijn voor patiënten en financiers. Patiëntenverenigingen willen die gegevens gebruiken om goede keuze-informatie te ontwikkelen voor cliënten en hun naasten<sup>12</sup>. Daarnaast wil het ministerie van VWS de benchmarkgegevens in de toekomst gaan gebruiken om bekostiging op te gaan baseren<sup>1</sup>.

***Gegevens hoeven niet 100% vergelijkbaar te zijn om ervan te kunnen leren.***

Dat de ROM-benchmarkgegevens zoveel doelen tegelijk moeten dienen is problematisch, omdat die doelen in de praktijk met elkaar botsen. Om te leren is bijvoorbeeld een omgeving nodig waarin mensen zich veilig voelen om van hun fouten te kunnen leren. In een afrekencultuur is die ruimte er niet. Op het niveau van benchmarken als intern leerinstrument lijken de methodologische problemen overkomelijk, juist omdat er geen harde conclusies aan worden verbonden, maar alleen een zoektocht naar mogelijke verklaringen voor verschillen. Kim de Jong: “Gegevens hoeven niet 100% vergelijkbaar te zijn om ervan te kunnen leren. Voor verantwoording is het wel problematischer. Daarom willen we dat er zorgvuldig met dit soort gegevens wordt omgegaan.” De huidige benchmarkgegevens op basis van ROM zijn ongeschikt als harde en enige maat van kwaliteit. Op het moment dat de data wel zo gebruikt gaan worden, zijn hogere methodologische eisen nodig.

***Je kan puur kijken naar of data volledig en bruikbaar zijn. Maar dat zegt natuurlijk niet alles. Je moet je ook afvragen hoe data in de praktijk verzameld worden.***

## Afrekenen met ROM?

Wanneer benchmarkgegevens worden gebruikt om op af te rekenen ontstaat nog een ander probleem waar veel mensen zich zorgen over maken: er kunnen perverse prikkels ontstaan die leiden tot bias of vertekening van de data. Roland Bal, hoogleraar Health Care Governance aan de Erasmus Universiteit (iBMG): “Je kan puur kijken naar of data volledig en bruikbaar zijn. Maar dat zegt natuurlijk niet alles. Je moet je ook afvragen hoe data in de praktijk verzameld worden. En hoe die metingen zelf het systeem



beïnvloeden. Aan die vragen wordt vaak voorbijgegaan (...) Wanneer derde partijen zoals zorgverzekeraars of de inspectie op basis van ROM-data gaan afrekenen gaan mensen die scores in de praktijk op een heel andere manier gebruiken. Er is dan in elk geval een risico dat mensen berekenend, strategisch met die ROM-scores omgaan. Dus dan gaat het regime de scores beïnvloeden. Als derde partijen het op zo'n manier gaan inzetten dan creëer je eigenlijk een pervers systeem.”

Ook psychiater Aartjan Beekman, hoofd van de afdeling psychiatrie van VUmc/GGz in Geest, maakt zich daar zorgen over. Hij licht dit toe aan de hand van een voorbeeld uit het onderwijs: “Ik woon in Hilversum, en dat gymnasium bij ons, daar slaagt 99% van de leerlingen. Daar zijn ze enorm trots en daar maken ze volop reclame mee. Maar ze zorgen er wel voor dat kinderen die een risico zijn, helemaal niet binnenkomen, of in de vierde afhaken. Dus ze zorgen ervoor dat de kinderen die uiteindelijk in de zesde uitkomen ook echt een heel grote kans hebben om te slagen. Dat doen ze door goed onderwijs te geven, maar ze doen het dus ook door selectie. En als je in de ggz de financiering af gaat laten hangen van de ROM-resultaten, dan krijg je ook in de zorg die perverse prikkel. En dan ga je natuurlijk zorgen dat je patiënten in zorg krijgt die goed opknappen. We kunnen dat als behandelaren best goed voorspellen. Dus dat zou een heel slechte ontwikkeling zijn.”

Het risico van perverse prikkels is reëel wanneer bekostiging op basis van ROM-data wordt doorgevoerd. Op dit moment worden behandelingen in de ggz door zorgverzekeraars nog op basis van diagnose-behandelcombinaties (dbc's) afgerekend. Hierdoor is nu de hoeveelheid en snelheid van geleverde zorg leidend bij de zorginkoop. Bij invoering van het stelsel van gereguleerde marktwerking in de ggz in 2008 was het uitgangspunt echter dat zorgaanbieders op kwaliteit zouden gaan concurreren. In de ggz wordt kwaliteit momenteel gebaseerd op drie verschillende uitkomsten; de effectiviteit van de zorg via ROM, de cliëntervaringen via



***We zien dat er door de inzet van vele partijen in het veld al veel bereikt is, maar dat er ook nog beperkingen zijn.***

de CQ-index en de veiligheid van de zorg via de Inspectie voor de Gezondheidszorg (IGZ) <sup>4</sup>. Wanneer uitkomstbekostiging wordt ingevoerd zouden ROM-gegevens een belangrijke rol gaan spelen bij het afrekenen van zorg. De Algemene Rekenkamer concludeerde eerder genoemde rapport echter dat de ROM-data nog niet van voldoende kwaliteit zijn om uitkomstbekostiging op te baseren <sup>1</sup>.

Maarten Erenstein is het daar niet mee eens: “Ons empirisch onderzoek naar de kwaliteit van de uitkomstinformatie van SBG is over het algemeen gunstig. Zo bleek uit recent onderzoek een goede betrouwbaarheid van de uitkomsten per zorgaanbieder in termen van stabiliteit over tijd.” Zowel SBG als ZN wijzen erop dat het rapport van de Algemene Rekenkamer niet berust op empirisch onderzoek, maar op een rondvraag onder sleutelfiguren in het veld. Nemen ze desondanks de conclusies serieus? Charlotte de Schepper, beleidsadviseur bij Zorgverzekeraars



Nederland (ZN): “De kritiek wordt door ZN in ieder geval gehoord. We zien dat er door de inzet van vele partijen in het veld al veel bereikt is, maar dat er ook nog beperkingen zijn. Daarom vinden verzekeraars het van belang om de benchmark nog verder door ontwikkelen, zodat kwaliteit nog beter inzichtelijk wordt.”

We vroegen een zorginkoper hoe zorgverzekeraars benchmarkgegevens momenteel gebruiken. Tomas Göbbels van VGZ geeft aan: “We verbinden geen consequenties aan behaalde delta T scores. We bespreken met aanbieders ROM-gegevens en vragen hoe we deze gegevens moeten interpreteren. Daar zijn we nog heel voorzichtig mee. Als VGZ willen we op de eerste plaats dat evaluaties plaatsvinden en dat ze ook daadwerkelijk gebruikt kunnen worden in de individuele behandeling. En op het moment dat dit gebeurt, hopen wij er uiteindelijk naartoe te gaan dat je instellingen met elkaar kunt gaan vergelijken, zodat we onze verzekerden naar een plek kunnen sturen waar de kans op herstel het grootst is.”

Zorgverzekeraars lijken nu nog terughoudend om harde consequenties aan ROM te verbinden. Ook omdat ze weten dat dit alleen maar tot weerstand zal leiden in het veld. Is er dan ook enige garantie dat zorgverzekeraars SBG data voorlopig niet voor de inkoop gaan gebruiken? Erenstein: “In de statuten van SBG staat opgenomen dat het doel is dat de uitkomstinformatie gebruikt moet gaan worden bij zorg in- en verkoop. Er zijn geen afspraken of garanties dat dit niet gebeurt. De uitkomstinformatie kan nu al een rol spelen bij de inkoop, maar dit moet wel op een creatieve en voor de professionals in de sector herkenbare manier worden ingezet. Plat bekostigen op uitkomst zal verschillen in behandeluitkomst en zorgkwaliteit tussen instellingen eerder groter maken in plaats van kleiner. Een betere weg zou zijn om gerichte afspraken te maken bij de zorginkoop over financiële beloning voor verbetering van de kwaliteit van zorg bij specifieke (diagnostische) doelgroepen in de ggz.”

In de onderhandelingen tussen zorginkopers en het professionele veld lijkt het nu juist aan die herkenbaarheid te ontbreken. Psychiater Beekman: “De zorgverzekeraars hebben de data en je merkt dat de getallen op de achtergrond al meespelen in de onderhandelingen over tarieven. Ik vind dat misbruik.” Omdat er nog geen officiële uitkomstbekostiging is ingevoerd, maar zorgverzekeraars wel al over globale uitkomstdata beschikken, is nu dus een situatie ontstaan waarin niet transparant is of en hoe zorgverzekeraars de ROM-data laten meewegen bij hun zorginkoop. Dat botst met het principe van een veilige leeromgeving die nodig is om succesvol te benchmarken. Het is dan ook niet vreemd dat juist over dit onderwerp overeenstemming is tussen zowel de sympathisanten van stopROM als de voorstanders van ROM: Afrekenen op basis van ROM uitkomsten is niet wenselijk.

### Hoe nu verder?

We hebben geconcludeerd dat benchmarken op basis van ROM-data vooral problematisch is doordat de cijfers voor verschillende doelen worden gebruikt. Leren, verantwoorden en afrekenen lopen in de praktijk door elkaar. Hierdoor ligt zowel strategisch gebruik van zorgaanbieders als misbruik van zorgverzekeraars op de loer. Hoe kunnen we die ontwikkelingen tegengaan? Daar zijn de meningen van de geïnterviewde over verdeeld. Beekman pleit voor het terugnemen van de regie: “Om iets tegen het misbruik van ROM data door zorgverzekeraars te doen kunnen doen, moet je zorgen dat je eerst zelf de baas bent over die data en dat je die data deelt als je denkt dat het goed is om die te delen, maar dat het niet automatisch gaat.” Toch is de vraag of we terug moeten willen naar een situatie waarin zorgaanbieders alleen op vrijwillige basis inzicht geven in hun resultaten. Cliënten vragen immers ook om informatie over de kwaliteit van de ggz<sup>12</sup>. De hamvraag is nu of er mogelijkheden zijn om kwaliteitsverbetering, verantwoording en zelfs uitkomstbekostiging op een goede manier samen te laten gaan. Roland Bal ziet meer heil in kwaliteitsverbetering door een verandering in gedrag, cultuur en houding,



dan in kwaliteitsbeoordeling waarbij de nadruk op regelgeving en systemen komt te liggen. Hij pleit dan ook voor een focus op benchmarken als leerinstrument. En: “Het is belangrijk dat dit in een veilige omgeving gebeurt, dus dat mensen niet het gevoel hebben dat ze worden afge-rekend, maar dat ze zich mogen verantwoorden, dat ze mogen leren.” Daarnaast raadt hij het professionele veld aan om voortdurend met externe partijen zoals de zorgverzekeraars en het ministerie in gesprek te blijven en te erkennen dat je met een gezamenlijk zoekproces bezig bent. Dat zoekproces kan echter niet geheel vrijblijvend zijn. “Ik denk dat je wel transparant moet zijn in een democratische samenleving. Iemand die in de publieke sector werkt, moet bereid zijn om uit te leggen wat hij doet en waarom. Maar je wilt je als zorgverlener wel verantwoorden over die zaken waarvan je denkt dat ze van belang zijn. En daar kun je ook wel over in gesprek gaan met patiënten, en verzekeraars om te definiëren wat het gemeenschappelijk belang is.”

***Je wilt je als  
zorgverlener wel  
verantwoorden over  
die zaken waarvan  
je denkt dat ze van  
belang zijn.***

Dat is precies wat ook Jan Hazelzet, hoogleraar waardegedreven zorg bij het Erasmus Medisch centrum ook nastreeft. Op basis van de wetenschappelijk literatuur en in samenspraak met cliënten en zorgverleners brengt hij in kaart wat per ziektebeeld de belangrijkste uitkomsten van een behandeling zijn (zowel medisch als niet-medisch), zodat die uitkomsten in de behandeling gemeten en gebruikt kunnen worden. Die uitkomsten worden in eerste instantie voor de

cliënt gemeten, maar Hazelzet heeft tegen het gebruik van die gegevens voor verantwoording geen bezwaar, integendeel: Hij hoopt dat deze uitkomsten de huidige ‘zinloze’ indicatoren gaan vervangen.

Nadat er in de afgelopen tien jaar grote systeemveranderingen in de zorg zijn doorgevoerd – met alle bijkomende administratieve lasten en een veel grotere rol van zorgverzekeraars – lijken de bezwaren van professionals hierop steeds meer gehoor te krijgen. In de onlangs verschenen zorgagenda<sup>13</sup> van de raad voor Volksgezondheid en Samenleving (RVS) is ‘verantwoording in dienst van verantwoordelijkheid’ een van de belangrijkste kernopgaves die de raad voor de toekomst ziet. Ze agendeert een zoektocht naar ‘nieuwe en andere vormen van verantwoording afleggen, die in eerste plaats als zinnig en behulpzaam wordt ervaren die eraan moeten voldoen’. ‘Die nieuwe vormen zullen gevolgen hebben voor de wijze van sturing door de overheid en controle door toezichthouders. Immers, niet de meetbare resultaten ten opzichte van het gemiddelde staan dan centraal, maar vooral de vraag of er geleerd en vernieuwd is.’

Die vraag zal ook in de verdere ontwikkeling van benchmarken in de ggz centraal blijven staan. Hoe zorgen we dat we benchmarken primair een leerinstrument blijft? Als beroepsgroep moeten we met elkaar, maar ook met cliënten en zorgverzekeraars in gesprek blijven over hoe we op een zinnige manier kwaliteit van behandelingen kunnen meten, waar ROM-data wel en niet geschikt voor zijn en hoe we perverse prikkels in het systeem kunnen beperken.



## Bronnen:

- <sup>1</sup> Algemene Rekenkamer. Bekostiging van de curatieve geestelijke gezondheidszorg (2017). [http://www.rekenkamer.nl/Publicaties/Onderzoeksrapporten/Introducties/2017/01/Bekostiging\\_van\\_de\\_curatieve\\_geestelijke\\_gezondheid](http://www.rekenkamer.nl/Publicaties/Onderzoeksrapporten/Introducties/2017/01/Bekostiging_van_de_curatieve_geestelijke_gezondheid)
- <sup>2</sup> <https://www.stopbenchmark.com/>
- <sup>3</sup> de Jong, K., Tiemens, B., Verbraak, M.J.P.M.e.a. (2017). Red ROM als kwaliteitsinstrument. Tijdschrift voor Psychiatrie, 59, 242-244. <http://www.tijdschriftvoorpsychiatrie.nl/assets/articles/59-2017-4-artikel-dejong.pdf>
- <sup>4</sup> van Os, J., & Delespaul, P. (2017) Een Kwaliteitskader voor de GGZ. <https://www.stopbenchmark.com/wp-content/uploads/2017/04/Kwaliteitskader-GGZ-2017-J.-van-Os.pdf>
- <sup>5</sup> Oudejans, S., Spits, M., & Schippers, G. (2011) Leren van uitkomsten - van taseki naar jiseki. Uit: van der Vlist, P., Mulder, N., van Hees, S.(red.) Van meten naar weten: ROM in de ggz. Amsterdam : Boom, 2011.
- <sup>6</sup> de Beurs, E., Barendregt, M., & Warmerdam, L (2017). Inleiding: een tipje van de sluier na vijf jaar benchmarken. Uit: Behandel uitkomsten: bron voor kwaliteitsbeleid in de ggz. Amsterdam : Boom, 2017.
- <sup>7</sup> Noom, R.M.J., de Jong, K., Tiemens, B., e.a. (2012). Routine outcome monitoring en benchmarking: hoe kunnen we behandelresultaten op een zorgvuldige manier vergelijken? Tijdschrift voor Psychiatrie, 54 (2), pp. 141 - 145.
- <sup>8</sup> Delespaul, A.E.G. (2015) Routine outcome measurement in the Netherlands-A focus on benchmarking. International Review of Psychiatry, Early Online: 1-9.
- <sup>9</sup> Stichting Benchmark GGZ. <https://www.sbggz.nl/Benchmarken?contentitem=e7f263cc-6c13-49fe-ae7c-a0e79daafd&paragraph=11d6d2e2-db10-4bd9-9110-aa9235635947#Confounding-van-uitkomsten-of-casemix-correctie>.
- <sup>10</sup> Nederlands Zorginstituut. <https://www.zorginzicht.nl/bibliotheek/paginas/transparantiekalender.aspx>.
- <sup>11</sup> van Venrooij, A.(2017). Kwaliteitsinformatie vanuit het perspectief van de patient. Uit: Behandeluitkomsten: bron voor kwaliteitsbeleid in de ggz.
- <sup>12</sup> MIND. <http://www.platformggz.nl/lpggz/newsitems/ni003332#.WN0PVzrnMIA.twitter>
- <sup>13</sup> Raad voor Volksgezondheid en Samenleving. <https://www.raadvv.nl/publicaties/item/de-zorgagenda-voor-een-gezonde-samenleving>

