

Benchmarken: doodlopende weg onder het mom van 'ROM'

J. VAN OS, J. BERKELAAR, A. HAFKENSCHIED E.A.^{1,2}



ARTIKEL



Er is grote zorg over de toenemende mate van 'bureaucrazy' in de zorg en de impact van deze doorgesloten vorm van bureaucratie op kwaliteit en werkbaarheid (Oosterhoff 2017). Damiaan Denys, voorzitter van de Nederlandse Vereniging voor Psychiatrie (NVvP), spreekt over 'een ver-slaving aan controle' in onze samenleving, die onder andere tot uiting komt in de uit de hand gelopen regeldruk in de zorg (van Kempen 2016). Hij heeft gelijk. GGZ Nederland (2017) kwam onlangs met een vernietigend rapport waarin staat dat psychiatrisch verpleegkundigen ruim 33% van hun behandelings-tijd kwijt zijn aan administratie en dat een gemiddelde GGZ-aanbieder zo'n 3500 A4'tjes (ruim 1 km papier) volschrijft om zich te verantwoorden naar de zorgverzekeraars.

Benchmarken in de ggz

Het benchmarkprogramma in de Nederlandse ggz spant op bureaucratisch gebied wellicht de kroon. Ggz-instellingen zijn verplicht om mee te doen met een nationale gestandaardiseerde dataverzameling voor de Stichting Benchmark GGZ (SBG), een stichting die bekostigd wordt door de verenigde zorgverzekeraars.

Het benchmarken is komen overwaaien uit de Verenigde Staten. In de vs geldt de onbewezen ideologie dat zinnige en zuinige zorg alleen kan gedijen als de gezondheidszorg wordt ingericht als een 'markt'. Het op grote schaal meten van vergelijkende benchmarkuitkomsten in de gezondheidsmarkt werd destijds gezien als een belangrijke innovatie. Eerst werd geëxperimenteerd met *pay-for-performance* (P4P) in de zorg. Later werd *hospital value-based purchasing* (VBP) ingevoerd. Volgens VBP krijgt een ziekenhuis alleen betaald als het 'value' kan aantonen, in de vorm van benchmarkuitkomsten.

Inmiddels weten we uit meerdere meta-analysen dat materieel belonen voor benchmarkkwaliteit volgens P4P niet werkt in de gezondheidszorg (Rosenthal & Frank 2006). Ook is er steeds meer wetenschappelijk bewijs dat het VBP-programma, oftewel het onderwerpen van de gezondheidszorg aan de tucht van de markt, niet werkt (Figuroa e.a. 2016; Papanicolas e.a. 2017).

Nederlandse politici met een voorkeur voor marktgericht denken omarmden de gedachte van beloning-op-basis-van-benchmarkuitkomsten. Zorgverzekeraars kregen een belangrijke rol bij de uitvoering. Zij riepen de SBG in het leven en dwongen alle Nederlandse ggz-aanbieders om hun patiënten te onderwerpen aan metingen in symptoomscores voor en na behandeling. Met deze metingen kon, in de logica van het marktdenken, de 'value' van de ggz bepaald worden, op basis waarvan VBP ingevoerd kon worden. De instelling met de meeste 'value' wordt beloond, degene met de minste 'value' valt uit de markt.

De Algemene Rekenkamer (2017) legde eerder dit jaar een bom onder het streven naar benchmarken in de ggz gericht op VBP. In haar rapport kwam de Algemene Rekenkamer (2017) tot de conclusie dat de data van SBG onvolledig, onbetrouwbaar en onvergelykbaar zijn en dus niet geschikt om te benchmarken. De kernhoogleraren psychiatrie waren eerder tot dezelfde conclusie gekomen (van Os e.a. 2012).

Het 'doorontwikkel'-argument

Onder voorstanders van VBP is een veelgehoord argument dat de SBG-dataverzameling 'doorontwikkeld' kan worden zodat benchmarken op termijn wél zinnig wordt. Het doorontwikkelargument, zoals ook aangegeven in het rapport van de Rekenkamer (2017), is feitelijk een ontkenning van de werkelijkheid. SBG probeert het benchmarken al 5 jaar te ontwikkelen, zonder ook maar in de buurt te komen van enige validiteit. Wat gaat nog eens 5 jaar, à raison van 30 miljoen euro directe en indirecte kosten per jaar, toevoegen, wanneer de literatuur nu al heeft uitgewezen dat VBP überhaupt niet werkt? Het meenemen van meer variabelen zoals leeftijd, geslacht en een soort postcode-indicator van sociaaleconomische status zal het casemix-probleem niet kunnen oplossen: de benchmark blijft altijd onzuiver. Alleen al de variatie tussen regio's is – om onbekende en daarmee niet corrigeerbare redenen – groter dan de variatie tussen instellingen.

Symptoomscores van traditionele vragenlijsten hebben een notoir zwakke relatie met wat patiënten in hun

bestaan van waarde vinden. Waarom zijn juist deze zeer beperkte instrumenten de kern geworden van het benchmarkargument? Bijvoorbeeld: wat is de benchmarkwaarde van een verschil van 0,1 standaarddeviatie op de Hamilton-depressieschaal?

Het belangrijkste probleem onder de 'benchmark' is onoplosbaar: risicoselectie aan de poort. De beste manier om de uitkomst van bijvoorbeeld prostaatcarcinoomchirurgie te beïnvloeden is om vooraf alleen die patiënten te selecteren bij wie succes verzekerd is. Zo ook in de ggz: een enge focus op selectie aan de poort, bijvoorbeeld gericht op burn-out of hoog functioneren in het autismespectrum, met uitsluiting van comorbiditeit van verslaving, persoonlijkheidsproblematiek, psychose, zelfverwaarlozing en suicidaliteit, garandeert de beste 'benchmark'-resultaten. Er is anekdotisch bewijs dat dit soort strategisch gedrag aan de orde van de dag is.

Strijdige belangen

De zorgverzekeraars hebben SBG mede opgezet om VBP te faciliteren. SBG heeft hiertoe een soort ranking van symptoomscoreveranderingen gemaakt. Er is toenemend anekdotisch bewijs dat zorgverzekeraars deze ranking gebruiken bij de zorginkoop. Op welke data en welke wetenschappelijke inzichten de SBG-rankings zijn gebaseerd, is volslagen onhelder. Hoewel er een hoogleraar ROM en Benchmarking is aangesteld met sponsoring door de SBG, is de 'ranking' niet *peer-reviewed*. De academische onafhankelijkheid staat bij voorbaat ter discussie. Onafhankelijke onderzoekers wordt tot nu toe, ondanks aanhoudende pogingen, toegang tot de SBG-data geweigerd. Binnen de ondoordringbare en ondoorzichtige SBG-vesting is een eigen variant van statistische benchmarking gecreëerd, op basis waarvan de zorgverzekeraars met de door hen gewenste VBP kunnen aanvangen.

Het 'ROM'-rookgordijn onder de SBG-benchmark

De voorstanders van benchmarken en VBP stellen de activiteiten van SBG in een gunstig daglicht. Ze leggen een verband tussen de dataverzameling van SBG en een respectabele klinische activiteit: *routine outcome monitoring* (ROM). Deze voorstanders doen het voorkomen alsof de activiteiten van SBG van belang zijn voor ROM. Onder het motto: SBG verzamelt ROM-data die in de klinische praktijk *sowieso* worden verzameld. Ergo: wie tegen SBG is, is tegen ROM en, in het verlengde daarvan, tegen *transparantie*. Feitelijk echter heeft de SBG-dataverzameling niets met ROM te maken. ROM is het aanbrengen van een flexibele *forward feedback loop* in de behandeling, op basis van persoonlijke behandeldoelen, zodat de behandeling voortdurend in gedeelde besluitvorming bijgestuurd kan worden (Hafkenscheid & Van Os 2014a; 2014b; van Os e.a. 2014).

AUTEURS

J. VAN OS, hoogleraar Psychiatrie, Maastricht UMC, Maastricht.

JUDICA BERKELAAR, patiëntenrechtenactiviste en ervaringsdeskundige.

ANTON HAFKENSCHIED, klinisch psycholoog-psychotherapeut, Arkin/Sinai Centrum, Joodse geestelijke gezondheidszorg.

WIEPKE CAHN, DIRK CORSTENS, PHILIPPE DELESPAUL, DAMIAAN DENYS, GERRIT GLAS, COBIE GROENENDIJK, PETER GROOT, MICHEL HENGVELD, WIM JONGEJAN, PETER VAN DER KORT, ELISABETH TER KULVE, ANNE MARSMAN, ERWIN VAN MEEKEREN, JOOST MERTENS, MENNO OOSTERHOFF, RUDOLF PONDS, ALAN RALSTON, FLOORTJE SCHEEPERS, REMKE VAN STAVEREN, AART SCHENE, JULES TIELENS, LEX VINK, SWANET WOLDHUIS.

CORRESPONDENTIEADRES

Prof. dr. Jim van Os, Universiteit Maastricht, afd. Psychiatrie en Psychologie, Euron, Postbus 616 (DRT 10), 6200 MD Maastricht.

E-mail: jvanos@maastrichtuniversity.nl

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 16-3-2017

TITLE IN ENGLISH

Benchmarking in mental health care: a dead end exercise disguised as 'ROM'

De SBG-dataverzameling is de antithese van ROM: een rigide gestandaardiseerde dataverzameling op basis van traditionele vragenlijsten die blind zijn voor persoonlijke behandeldoelen (Hafkenscheid & van Os 2013; 2016). Uit de literatuur is genoegzaam bekend dat dit soort rigide dataverzamelingen het klinische proces verstoren en elke bruikbaarheid in de praktijk missen (Cofiel e.a. 2013). Weinig klinici in Nederland onderschrijven de bewering dat de SBG-dataverzameling iets te maken heeft met ROM. Integendeel: de afgedwongen dataverzameling voor de SBG lijkt de klinische ROM grotendeels te hebben verdrongen. In de praktijk worden de SBG-data verzameld op een manier die compleet losgezongen is van de klinische praktijk: via websites en callcenters en in de wachtkamers vult de patiënt lijstjes in, die nooit in de spreekkamer gebruikt zullen worden. Behandelaars die hun eigen klinische ROM niet moedeloos vaarwel hebben gezegd, nemen hun instrumenten af *ondanks*, en niet *dankzij* de SBG-dataverzameling. De SBG-dataverzameling blinkt vooral uit in het

verzwaren van de bureaucratische belasting binnen de ggz. De argumenten om het systeem te redden doen steeds gekunstelder en geforceerder aan.

De olifant in de kamer: grootschalige privacy-schending

Wellicht de grootste olifant in de benchmarkkamer is de privacy-schending. In de ggz is de implementatie van de SBG-dataverzameling vrij geruisloos gegaan, zonder noemenswaardige privacyoverwegingen. Recent echter werd de praktijk bij de Nederlandse Federatie van Universitair Medische Centra tegen het licht gehouden. De conclusie: er zijn serieuze privacyproblemen. Ten eerste wordt patiënten een wettelijk verplichte inspanning opgelegd voor een foutief inkoopbeleid door de zorgverzekeraars. Gezien het verplichte karakter kan er geen sprake zijn van vrijwillig verleende toestemming. Ten tweede kunnen vraagtekens geplaatst worden bij de belangeloosheid van SBG. De stichting wordt immers bekostigd door de zorgverzekeraars.

Hoeveel patiënten zijn zich ervan bewust dat hun data in gepseudonimiseerde vorm (dat wil zeggen: koppelbaar, in tegenstelling tot geanonimiseerd) over verslaving, seksualiteit, suicidaliteit en paranoia naar een stichting van de zorgverzekeraar worden gestuurd, waar ze *in principe* met andere data kunnen worden gekoppeld? En wie wil en moet ze dat vertellen? De zorginstelling, de primaire behandelaar, de zorgverzekeraar? Geen van drieën heeft daar enig belang bij.

In juridische termen voldoen ggz-instellingen niet aan het criterium van *good care* wanneer ze de data van hun patiënten in koppelbare vorm bij een stichting van de zorgverzekeraar terecht laten komen. De ggz staat onder dermate zware economische en juridische druk om de SBG data te leveren dat deze druk steeds meer wordt doorvertaald naar de patiënten. Via websites van instellingen wordt patiënten verteld dat de instelling inkomsten misloopt als de patiënten geen vragenlijsten invullen. Sommige ggz-instellingen roepen ROM-callcenters in het leven waar patiënten door studenten telefonisch worden overvallen om mee te gaan doen met benchmarken. Patiënten krijgen te horen dat de vragenlijsten *'van belang zijn voor de behandeling'* (anders gezegd: onderdeel zijn van 'ROM'), terwijl ze feitelijk in de behandeling geen enkele rol spelen.

In de General Data Protection Regulation (GDPR) is vastgelegd welke privacyregels vanaf 2018 gelden binnen de hele EU. Het staat nu al vast dat de huidige SBG-benchmarkprocedure de GDPR-toets niet kan doorstaan. De GDPR eist bijvoorbeeld dat er een duidelijk en realiseerbaar doel ten grondslag ligt aan de verwerkte persoonsgegevens (dus geen *'benchmark in progress'*) en er ondubbelzinnige toestemming is van de betrokken burger om de gegevens te

verwerken, die bovendien op elk moment ingetrokken kan worden.

Conclusie

Gevoelige data van naar schatting 500.000 Nederlanders liggen, zonder dat zij hiervan op de hoogte zijn en zonder dat hun privacy met 'harde' waarborgen is beschermd, bij een stichting van de zorgverzekeraars. Daar worden de data zonder peer review aan niet-transparante statistische analyses blootgesteld. Zorgverzekeraars menen daarmee iets in handen te hebben dat voor een 'benchmark' door kan gaan, op basis van de achterhaalde gedachte dat vpb mogelijk en wenselijk is. Wij beschouwen dit als een ondoorzichtige, niet-valide onderneming, waarmee we jaarlijks 30 miljoen euro over de balk gooien.

Met dit gebeuren zijn belangen gemoeid van diverse 3- tot 5-letterige peri-ggz-partijen, alsmede partijen die zich strategisch rond de SBG-dataverzameling hebben georganiseerd. Dit zijn in de eerste plaats financiële belangen, maar ook eer en het voorkomen van gezichtsverlies. Voor GGZ Nederland, de brancheorganisatie, is het bijvoorbeeld moeilijk om afstand te nemen van een praktijk waarmee zij ooit zelf akkoord is gegaan. En hetzelfde geldt voor de diverse beroepsverenigingen, die krampachtig vasthouden aan het 'argument', dat de 'SBG-ROM' *'work in progress'* zou zijn.

Hoe nu verder? De ggz heeft behoefte aan een nieuw, valide en werkbaar kwaliteitskader dat zich durft te emanciperen ten opzichte van achterhaald marktdenken. In een aankomend nummer van het *Tijdschrift voor Psychiatrie* zullen wij uitgebreid ingaan op een nieuw kwaliteitskader voor de ggz, uitgaande van de intrinsieke motivatie van hulpverleners om goede zorg te leveren en daarover verantwoording af te leggen.

NOOT

1. Namens alle in de adresnoot genoemde auteurs.
2. Voor alle auteurs geldt dat de bijdrage op persoonlijke titel is geleverd en dus niet gelezen mag worden als het officiële standpunt van de instellingen waar zij werkzaam zijn.

LITERATUUR

- Algemene Rekenkamer. Bekostiging van de curatieve geestelijke gezondheidszorg. Den Haag: Algemene Rekenkamer; 2017.
- Cofiel L, Bassi DU, Ray RK, Pietrobon R, Brentani H. Detecting dissonance in clinical and research workflow for translational psychiatric registries. *PLoS One*. 2013; 8: e75167.
- Figueroa JF, Tsugawa Y, Zheng J, Orav EJ, Jha AK. Association between the value-based purchasing pay for performance program and patient mortality in US hospitals: observational study. *BMJ* 2016; 353: i2214.

- GGZ Nederland. Het roer moet om: Onderzoek naar de administratieve lasten in de geestelijke gezondheidszorg. Amersfoort: GGZ Nederland; 2017.
- Hafkenscheid A, van Os J. Huidige ROM doet afbreuk aan valide kwaliteitsmeting. *Tijdschr Psychiatr* 2013; 55: 179-81.
- Hafkenscheid A, van Os J. Naar een deugdelijke ROM. *Maandblad Geestelijke Volksgezondheid* 2014a; 69: 20-8.
- Hafkenscheid A, Van Os J. ROM van geïndividualiseerde behandeldoelen. *Psychopraktijk* 2014b; 6: 29-32.
- Hafkenscheid A, van Os J. Wat ieder die betrokken is bij ROM zich over de metingen moet realiseren. *Tijdschr Psychiatr* 2016; 58: 388-96.
- Kempen J van, Damiaan Denys: 'Psychiaters bereid op de barricaden te gaan'. *Het Parool* 15-4-2016.
- Os J van, Delespaul P, Barge D, Bakker RP. Testing an mHealth momentary assessment Routine Outcome Monitoring application: a focus on restoration of daily life positive mood states. *PLoS One* 2014; 9: e115254.
- Os J van, Kahn R, Denys D, Schoevers RA, Beekman AT, Hoogendijk WJ, e.a. ROM: gedragsnorm of dwangmaatregel? Overwegingen bij het themanummer over routine outcome monitoring. *Tijdschr Psychiatr* 2012; 54: 245-53.
- Oosterhoff M. Bureaucrazy: Weblog, *Medisch Contact*; 27-1-2017.
- Papanicolas I, Figueroa JF, Orav EJ, Jha AK. Patient hospital experience improved modestly, but no evidence medicare incentives promoted meaningful gains. *Health Aff (Millwood)* 2017; 36: 133-40.
- Rosenthal MB, Frank RG. What is the empirical basis for paying for quality in health care? *Med Care Res Rev* 2006; 63: 135-57.