

Achtergronden en gebruiksmogelijkheden van ROM in de ggz

M.A. NUGTER, V.J.A. BUWALDA

ACHTERGROND In dit artikel beschrijven wij de ontwikkelingen op het gebied van kwaliteitszorg en wetenschappelijk onderzoek die hebben geresulteerd in een wereldwijde aandacht voor routine outcome monitoring (ROM).

DOEL Inzicht geven in de betekenis van ROM voor de Nederlandse ggz.

METHODE Gebruik is gemaakt van wetenschappelijke literatuur over onderzoek, kwaliteit, meetinstrumenten en implementatie.

RESULTATEN ROM betekent het routinematig gebruiken van meetinstrumenten om uitkomsten van behandelingen te meten, ook tussentijds. Dit kan worden beschouwd als de checkfase van de 'plan-do-check-act'-cyclus. Inhoudelijke, wetenschappelijke en praktische criteria liggen ten grondslag aan de keuze van meetinstrumenten. ROM kan op vier manieren worden ingezet: voor de behandeling van individuele patiënten, om beleid te toetsen, voor benchmarking en voor onderzoek. Voor de implementatie is een gecombineerde top-down-, bottom-upbenadering te prefereren boven hetzij een top-down-, hetzij een bottom-upaanpak.

CONCLUSIE Door de systematiek en de aggregatiemogelijkheden kan ROM de ggz veel bieden.

[TIJDSCHRIFT VOOR PSYCHIATRIE 54(2012)2, 111-120]

TREFWOORDEN behandelplannen, benchmarking, effectiviteitsonderzoek, implementatie, PDCA-cyclus, routine outcome monitoring /measurement (ROM)

Systematisch meten van uitkomsten is in de ggz de laatste jaren enorm in de belangstelling komen te staan. Aanvankelijk was het meten van uitkomsten onderdeel van kwaliteitsmanagement (Nugter e.a. 1998; Walburg & Brinkmann 2001); later bleek ook dat uitkomsten meten de effectiviteit van behandelingen kan verhogen (Lambert 2007; Miller e.a. 2005; Slade 2002). Echt brede belangstelling ontstond na de eerste ervaringen met de prestatie-indicatoren en door de toenemende invloed van zorgverzekeraars, die bij de zorginkoop de kwaliteit willen betrekken. Inmiddels zijn voor het systematische gebruik van meetinstrumenten om uitkomsten te meten de termen 'routine out-

come monitoring' (De Beurs & Zitman 2007) of 'routine outcome measurement' (ROM) gangbaar geworden.

In dit inleidend artikel gaan we in op de vraag wat ROM inhoudt en welke mogelijkheden deze biedt om de kwaliteit en effectiviteit van de ggz te toetsen en te verbeteren.

VOORGESCHIEDENIS VAN ROM

Onderzoek en uitkomsten

De wetenschappelijke belangstelling voor uitkomsten van psychiatrische behandelingen in

de alledaagse praktijk is in belangrijke mate te danken aan het toenemend besef dat de resultaten van gerandomiseerde gecontroleerde trials (RCT's) voor de praktijk beperkt toepasbaar zijn. Deze beperkingen zijn inherent aan de strenge experimentele condities waaronder dergelijk onderzoek moet plaatsvinden: de strenge in- en exclusiecriteria, behandelingen die in tijd gelimiteerd zijn, volgens een strak behandelprotocol met het oog op de *treatment fidelity*, en grote testbatterijen die voor sommige patiëntengroepen een drempel zijn voor deelname (Nugter 2000).

Deze strenge condities zijn nodig voor de interne validiteit, maar ze gaan ten koste van de externe validiteit. Immers, in de praktijk voldoen grote groepen patiënten niet aan de in- en exclusiecriteria, vanwege leeftijd of comorbiditeit, en ook wijken behandelingen in de praktijk vaak af van de onderzochte. Dat geldt voor medicamenteuze behandelingen, waarbij bijvoorbeeld voorschriften voor dosis en tijdstippen door artsen en patiënten niet altijd kunnen worden opgevolgd, maar ook voor psychotherapie.

Voor langdurende behandelingen zijn RCT's moeilijk uitvoerbaar, vanwege ethische bezwaren om patiënten lang in een controleconditie te houden en omdat het lastig is hen zo lang gemotiveerd te houden. Cohortonderzoek is dan wetenschappelijk het hoogst haalbare. Hierbij relateert men de resultaten van behandelingen in de praktijk aan variabelen die deze resultaten kunnen beïnvloeden, zoals leeftijd, geslacht, comorbiditeit en duur van de klachten. In de Engelstalige literatuur spreekt men van onderzoek naar *effectiveness* (doelmatigheid), tegenover onderzoek naar *efficacy* (werkzaamheid; RCT).

Een extra impuls voor de wetenschappelijke aandacht voor de resultaten van behandelingen in de dagelijkse praktijk is het toenemende bewijs dat deze minder positief zijn dan die in RCT's. Dit rechtvaardigt het toetsen van wetenschappelijke bevindingen in de praktijk (Hansen e.a. 2002; Lambert & Ogles 2004).

Kwaliteit en uitkomsten

Donabedian (1966) maakte in de jaren zestig onderscheid tussen structuur-, proces- en uitkomstaspecten van kwaliteit. Structuuraspecten betreffen de voorwaarden waaronder behandelingen worden geboden. Procesaspecten betreffen de wijze waarop dat gebeurt en uitkomsten betreffen zowel klinische als bedrijfsmatige resultaten. Veel aandacht werd aanvankelijk besteed aan de proces- en de structuuraspecten van kwaliteit, vanuit de gedachte dat het met de resultaten vanzelf goed zou komen als de processen maar in orde waren.

In de jaren tachtig kwam daar in de vs, door de opkomst van de *managed care* verandering in. *Managed care* richt zich op methoden die tot doel hebben de zorg op economische wijze in te richten. Kenmerkend is dat informatiemanagement door een derde partij wordt verzorgd, dus niet door de arts en/of de patiënt (Hooghiemstra 1998).

Ellwood (1988) introduceerde de term '*outcomes management*'. Hij bepleitte een database met informatie en analyses over klinische, financiële en gezondheidsuitkomsten die voor patiënten begrijpelijk zijn. Daarmee zou men de relaties tussen medische interventies en de gezondheidsuitkomsten, en tussen gezondheidsuitkomsten en kosten beter kunnen schatten en zou men keuzes in de zorg meer op feiten kunnen baseren.

Meten van uitkomsten als aspect van de kwaliteit van zorg lijkt op *effectiveness* onderzoek, maar is niet hetzelfde, omdat het wordt ingezet voor continue kwaliteitsverbetering (Dickey 2005; Hermann 2005). Dit betekent dat de gegevensverzameling en -verwerking meer in een praktische dan in een wetenschappelijke context plaatsvinden. Aan de meetinstrumenten worden behalve wetenschappelijke (psychometrische) ook praktische eisen gesteld (zie verderop). Over de resultaten wordt niet zozeer gerapporteerd aan een wetenschappelijk forum, maar aan eigen zorgprofessionals en patiënten van de ggz. Dat stelt eisen aan de wijze en het moment waarop dat gebeurt, wie dat doet en wie ervoor verantwoordelijk is dat de juiste informatie bij de juiste personen terechtkomt (Dickey 2005).

Uitkomsten meten als onderdeel van de kwaliteitscyclus

Wetenschappelijke kennis komt tot stand door het systematisch doorlopen van de empirische cyclus (De Groot 1961; Kerlinger 1986): het formuleren van een hypothese (en nulhypothese) die door inductie en deductie tot stand is gekomen, het bedenken en uitvoeren van een methode om de hypothese empirisch te toetsen (de nulhypothese te verwerpen) en vervolgens met inductie op basis van de gegevens conclusies trekken over de hypothese.

Deze empirische cyclus vinden we ook terug in de cyclus van Deming (1982), die de basis is voor het werken aan kwaliteitsverbetering. Deze PDCA-cyclus bestaat uit vier stappen: plannen (*plan*), uitvoeren (*do*), toetsen (*check*) en verbeteren (*act*). Organisaties plannen hun activiteiten in het verlengde van hun expliciete doelstellingen (*plan en do*). De uitgevoerde activiteiten worden getoetst aan de hand van tevoren overeengekomen kwaliteitskenmerken en -normen (*check*). De evaluatie van de resultaten leidt tot het bijstellen van de uitvoering of, in voorkomende gevallen, van de gestelde doelen of criteria (*act en plan*).

Dit continue proces van kwaliteitsverbetering krijgt binnen ROM op verschillende niveaus (micro-, meso- en macroniveau) zijn beslag. Op geaggregeerd niveau zijn de ROM-gegevens onderdeel van de fase waarin getoetst wordt in hoeverre een afdeling of organisatie de doelstellingen heeft behaald. Het opstellen, uitvoeren, evalueren en bijstellen van behandelplannen voor patiënten vormen feitelijk een vereenvoudigde empirische cyclus. Scores op meetinstrumenten aan het begin van de behandeling kunnen uitgangspunt zijn voor de doelen van het behandelplan. De tussentijdse metingen en eindmetingen zijn de toetsfase, waarin men vaststelt in hoeverre de doelen worden gerealiseerd tijdens het behandelproces. Daarop kan men vervolgacties inzetten: doorgaan met de afgesproken behandeling, het behandelplan bijstellen of de behandeling beëindigen.

Het verzamelen van gegevens over uitkomsten heeft alleen zin als er mogelijke verklaringen zijn voor die uitkomsten, als vanuit de toetsfase kan worden teruggegrepen naar het doel (plannen) en naar de activiteiten die zijn ondernomen (uitvoeren). Bij uitkomsten van behandelingen zoekt men die verklaringen in de geboden behandeling, het type patiënten aan wie de behandeling geboden wordt, de behandelaar die de behandeling aanbiedt en de omstandigheden waaronder deze dat doet. Het systematisch vaststellen van uitkomsten heeft alleen zin als koppeling met processen en structuuraspecten en met patiëntkenmerken (casemix) plaatsvindt. Dat geldt ook voor het vergelijken van uitkomsten. In hun artikel over vergelijkbaarheid gaan Noom e.a. elders in dit nummer daar verder op in.

Routinematig uitkomsten meten

Aandacht voor het routinematig meten van behandeluitkomsten was er aanvankelijk vooral in de VS en het Verenigd Koninkrijk. Daar werden ook instrumenten ontwikkeld om uitkomsten te meten. De *Outcome Questionnaire* (OQ; Lambert e.a. 2004), de *Health of the Nation Outcome Scales* (HONOS; Wing e.a. 1998) en de *Outcome Rating Scale* (ORS; Miller e.a. 2003) zijn bekende voorbeelden (zie ook Buwalda e.a. dit nummer). In de vorige eeuw was het Standaard Evaluatie Project (STEP) in de klinische psychotherapie een van de weinige projecten in Nederland waarin uitkomsten van verschillende instellingen routinematig werden vergeleken met oog op kwaliteitverbetering (Nugter e.a. 1998).

Meer recent zijn er in diverse sectoren van de ggz projecten gestart die tot doel hebben systematisch uitkomsten te meten en is serieus werk gemaakt van psychometrisch onderzoek naar de kwaliteit en de bruikbaarheid van uitkomstinstrumenten voor de Nederlandstalige populatie (De Jong e.a. 2007; Hafkenscheid 2010a; Mulder e.a. 2004). Van meet af aan ging het niet alleen om het meten aan het begin en aan het einde, maar ook tussentijds, waarmee het ook mogelijk werd om

de resultaten te gebruiken voor de individuele behandeling.

Onderzoek in de vs had inmiddels uitgewezen dat feedback aan de behandelaar over het scoreverloop de effectiviteit en daarmee de kwaliteit van de behandeling verhoogt (Lambert e.a. 2001). Deze verbetering van de effectiviteit en de kwaliteit zou ook de enige rechtvaardiging voor ROM moeten zijn. Inmiddels lijken vooral behandelingen van patiënten die (te) weinig vooruitgang boeken te profiteren van de toevoeging van ROM (zie artikel van Carlier e.a. elders in dit nummer).

De motivatie om ROM te gebruiken neemt toe als de behandelaar de scores van de instrumenten direct kan gebruiken om individuele behandelingen zo nodig bij te stellen (De Jong & Nugter 2011). En als de behandelaar dat doet, dan zijn ook patiënten eerder bereid vragenlijsten in te vullen. Dat betekent dat voor de inrichting van een systeem van ROM en voor de implementatie ervan, patiënten en behandelaars centraal moeten staan.

KEUZE VAN MEETINSTRUMENTEN

Bij de keuze van meetinstrumenten voor ROM moeten we een onderscheid maken tussen generieke meetinstrumenten, die voor uiteenlopende groepen patiënten kunnen worden gebruikt, en stoornisspecifieke instrumenten. Wiger en Solberg (2001) raden een efficiënte combinatie van beide aan. Generieke instrumenten maken het mogelijk om op groepsniveau de resultaten vast te stellen en om het individuele beloop van scores af te zetten tegen het beloop van scores op groepsniveau. Stoornisspecifieke instrumenten zijn van belang om verbeteringen in specifieke symptomen bij diagnostisch homogenere subgroepen patiënten te kunnen volgen, bijvoorbeeld bij het toepassen van de richtlijn voor depressie.

De criteria voor de keuze van de instrumenten vallen uiteen in inhoudelijke, wetenschappelijke en praktische. De wetenschappelijke criteria komen aan de orde in het artikel van Buwalda e.a. in dit nummer.

Inhoudelijke criteria

Bij de keuze van de meetinstrumenten is het belangrijkste criterium de klinische relevantie: meet het instrument de gebieden waar het in de behandeling om gaat? Denk bijvoorbeeld aan: afname van klachten, verbeterd functioneren, meer kwaliteit van leven. De keuze van meetinstrumenten verschilt per doelgroep. Bij patiënten die relatief kort worden behandeld voor angst of depressie zal het gaan om de afname van hun angstige of depressieve klachten en bijvoorbeeld hun sociaal-maatschappelijk functioneren dat door hun klachten negatief kan worden beïnvloed. Bij patiënten die langdurend van zorg afhankelijk zijn, zal het gaan om de ernst van de klachten, om hun kwaliteit van leven en de beperkingen in de verschillende levensgebieden. In dit themanummer gaan Zitman e.a., Boer, Van Marle e.a. en Oudejans e.a. hier specifieker op in.

Een ander belangrijk inhoudelijk aspect betreft de keuze voor zelfinvullijsten of beoordelingsinstrumenten. Veel patiënten die ambulantly worden behandeld of begeleid, kunnen goed zelf rapporteren over de ernst van hun problematiek of hun kwaliteit van leven, al moet men wel rekening houden met de mogelijkheid van sociaal wenselijke antwoorden. Voor patiënten die erg in de war zijn, zijn zelfinvullijsten niet geschikt en moet men uitwijken naar beoordelingslijsten. Voor het vaststellen van psychotische verschijnselen en voor risicotaxaties zijn er alleen beoordelingslijsten.

Praktische criteria

ROM is alleen zinvol wanneer meetinstrumenten in de loop van de behandeling herhaald worden afgenomen, omdat we alleen dan verandering kunnen vaststellen. Dat betekent dat we ook praktische eisen moeten stellen aan de meetinstrumenten, in termen van tijd, geld en gebruiksgemak (Dickey 2005). Een beperkte testbatterij zal gemakkelijker worden geïmplementeerd dan een uitgebreide. Meetinstrumenten die speciaal voor

ROM worden ontwikkeld, zijn vaak ook kort; de ORS en de *Session Rating Scale* (SRS; Hafkenscheid 2010a) hebben elk zelfs slechts vier items.

Patiënten blijken over het algemeen bereid om vragenlijsten in te vullen, zeker als dit voor de behandeling is, maar bij herhaald afnemen is uitval vaak aan de orde (zie Mulder e.a. elders in dit nummer). Het is daarom belangrijk om meetinstrumenten in te zetten die eenvoudig en snel zijn in te vullen.

Ook voor de professional is het van belang dat er niet te veel meetinstrumenten worden gebruikt. Deze heeft beperkte tijd en moet de informatie tot zich kunnen nemen. Hoe meer meetinstrumenten, hoe meer tijd het kost, des te groter de kans dat de informatie verkregen met het gebruik van de meetinstrumenten lastig te interpreteren is – en des te groter de kans dat die informatie niet wordt gebruikt (De Jong & Nugter 2011). Dit speelt zeker bij een grote caseload.

Omdat het bij ROM over grote aantallen patiënten en afnames gaat, hebben gratis meetinstrumenten de voorkeur, mits ze voldoen aan de inhoudelijk en psychometrische criteria.

GEBRUIK VAN ROM

Gebruik op individueel niveau

Er zijn meerdere manieren om scores op meetinstrumenten te gebruiken (zie ook Buwalda e.a. 2011). Allereerst kan men meetinstrumenten periodiek afnemen, bijvoorbeeld bij het opstellen en evalueren van behandelplannen (Wiger & Solberg 2001). Bij iedere volgende evaluatie van het behandelplan kan men de afnames herhalen, zodat het beloop van de scores meebepalend is voor beslissingen over de verdere behandeling. Scores op meetinstrumenten dienen dan bij voorkeur onderdeel te zijn van het behandelplan en kunnen ook worden gebruikt om behandeldoelen concreet te formuleren. Dit kan bijvoorbeeld door aan te geven hoeveel punten een patiënt op een meetinstrument vooruit moet zijn gegaan (Wiger & Solberg 2001). Deze koppeling met de evaluatie

van het behandelplan en met de behandeldoelen is in principe in iedere setting mogelijk.

Een tweede manier om meetinstrumenten te gebruiken is om beslissingen over tussentijdse wijzigingen in de behandeling te ondersteunen en de uitkomsten te evalueren. Als men bijvoorbeeld overweegt om de behandeling van een patiënt (tijdelijk) te intensiveren met een opname kan men scores op meetinstrumenten gebruiken om te beslissen of dat nodig is. Of als een patiënt op basis van de richtlijnen medicatie krijgt voorgeschreven, dan kan het scorebeloop laten zien of het effect binnen de voorgeschreven periode ook bereikt is.

Een derde manier om meetinstrumenten op individueel niveau te gebruiken wordt ‘tracking’ genoemd. Dat houdt in: iedere sessie meten. Soms kiest men voor tracking omdat de implementatie ervan relatief gemakkelijk is: men hoeft niet na te denken over de momenten waarop er gemeten moet worden en de koppeling daarvan met het behandelproces. Tracking lijkt vooral aangewezen voor patiënten bij wie veranderingen op korte termijn te verwachten zijn.

Bij individueel gebruik is ook de vraag van belang of en hoe de informatie gedeeld moet worden met patiënten. De hulpverlener dient zich goed te realiseren wat de consequenties zijn van het delen van informatie met zijn patiënt, ook al lijkt het vanzelfsprekend om dit altijd te doen. De eerste consequentie is dat de informatie voor patiënten inzichtelijk moet zijn met daarbij een duidelijke conclusie (u gaat vooruit, u blijft stabiel of u gaat achteruit). Daarnaast moet voor patiënten duidelijk zijn dat de scores niet de enige bron van informatie zijn voor de mate van verandering. Eigen indrukken van patiënten over andere voor hen belangrijke (niet gemeten) aspecten van de behandeling en de klinische deskundigheid van de behandelaar zijn minstens zo relevant. We moeten voorkomen dat de behandelaar en zijn of haar patiënten zich alleen nog focussen op het omlaag brengen van de scores op de meetinstrumenten.

Een andere factor is hoe behandelaars aankijken tegen de feedback die met de meetinstrumenten

ten wordt verkregen. Daarbij spelen de kenmerken van de feedback en de persoonlijkheidsaspecten van de behandelaar een rol (De Jong & Nugter 2011). De Jong en Hafkenscheid gaan beiden elders in dit nummer verder in op de ervaring van de behandelaar met feedback.

Gebruik op geaggregeerd niveau binnen instellingen

Als er van een substantieel deel van de patiënten gegevens zijn, wordt het interessant om uitkomsten van verschillende onderdelen binnen een organisatie te analyseren en te presenteren aan de professionals. De ervaringen met het STEP-project (Nugter e.a. 1998) en met de implementatie van ROM in verschillende settings (Buwalda e.a. 2011; Nugter 2008) leren tot nu toe dat terugkoppeling positief wordt ontvangen. Uiteraard is het van belang om niet alleen uitkomsten te presenteren, maar ook te analyseren of en hoe de uitkomsten gerelateerd zijn aan patiëntkenmerken en de geboden zorg (Hermann 2005). Daarbij is wetenschappelijke literatuur over voorspellers van behandelresultaten nuttig, evenals de ervaring van professionals. Het is ook mogelijk om meetinstrumenten die het behandelproces meten onderdeel te maken van de ROM, zoals in het Leuven Study Case Protocol gebeurt (zie de bijdrage van Stinckens elders in dit nummer) en de *Client Directed Outcome Informed*-benadering (CDOI; Hafkenscheid 2010b).

Het ligt voor de hand om vergelijkingen te maken tussen afdelingen, teams of patiëntengroepen (Hermann 2005; Nugter 2008). Ook kan men met herhaalde metingen nieuw beleid evalueren. Dit hoeven geen zuiver wetenschappelijke onderzoeksexercities te zijn. Het gaat erom de behandelaars de verzamelde gegevens terug te geven met als doel ervan te leren, erover te reflecteren en beleid te toetsen.

Gebruik op geaggregeerd niveau: vergelijkingen tussen instellingen

In 2009 is door de koepelorganisatie GGZ Nederland het project 'ROM GGZ' gestart. Doel is om in 2014 bij 50% van de patiënten in de ggz routinematig gebruik te maken van meetinstrumenten. In het project besteedt men veel aandacht aan het synchroniseren van de te gebruiken meetinstrumenten en aan methoden om aggregatie van gegevens vlot te laten verlopen en om benchmarking mogelijk te maken. Samengewerkt wordt met Zorgverzekeraars Nederland. Voor de keuze van meetinstrumenten en definiëren van uitkomsten zijn een 'werkgroep vergelijkbaarheid' en expertgroepen in het leven geroepen.

Bij benchmarking vergelijkt men de eigen uitkomsten met die van anderen of met een standaard. Door vergelijking kan men uitkomsten beter interpreteren. Uiteraard moet men daarbij dan rekening houden met de factoren die van invloed zijn op de vergelijkbaarheid (zie de bijdrage van Noom e.a. in dit nummer en Delespaul 2010). Het gebruik van gegevens voor externe doeleinden kan interfereren met interne doelen. Zo kan het idee dat gegevens 'naar buiten gaan' een defensieve reactie bij professionals tot gevolg hebben, wat de implementatie kan bemoeilijken. Los daarvan kan de wetenschap dat deze gegevens ook worden gebruikt voor externe verantwoording van invloed zijn op de manier waarop professionals en patiënten de meetinstrumenten invullen. Het is goed om zich dat te realiseren.

Gebruik voor onderzoek

Goedwerkende ROM-systemen leveren een schat aan gegevens en kunnen nieuwe inzichten genereren. Onderzoek met ROM is veelal gericht op effectiviteit van behandelingen zoals die in de alledaagse klinische praktijk plaatsvinden.

Sytema e.a. (2011) beschrijven voorwaarden waaraan moet zijn voldaan om ROM-gegevens te kunnen gebruiken voor onderzoek. De belangrijkste daarvan betreffen de psychometrische kwali-

teiten van de te gebruiken instrumenten, het zoveel mogelijk voorkomen van bias bij beoordelingen of van bias door selectieve uitval (zie ook Mulder e.a. in dit nummer), controle op de interbeoordelaarsbetrouwbaarheid en de aanwezigheid van gegevens over patiëntkenmerken (case-mix). Bij evaluaties van de effectiviteit van behandelingen dient men die behandelingen zelf goed te beschrijven. Dit om te voorkomen dat het niet duidelijk is wat men evalueert, waardoor de bevindingen niet overdraagbaar zijn.

In Nederland lopen op dit moment diverse onderzoeken naar de effectiviteit van ROM als werkmodel: aan de Erasmus Universiteit het project Monitoring Psychotherapie (www.mppsychotherapie.nl) en bij Altrecht (Buwalda e.a.) en bij GGZ Noord-Holland-Noord (De Jong e.a.; www.ggz-nhn.nl) lopen projecten. Het gaat in alle gevallen om onderzoek naar de effectiviteit van ROM bij patiënten die zich hebben aangemeld voor een kortdurende ambulante behandeling. Al deze projecten zijn inmiddels in een afrondende fase en naar verwachting verschijnen in de loop van 2012 publicaties over de resultaten.

IMPLEMENTATIE

Buwalda e.a. (2011) geven voorbeelden van implementatieprojecten in de langdurende psychiatrie (GGZ Noord-Holland-Noord, Delta Psychiatrische Centrum), bij poliklinische behandelingen en in de klinische setting (Altrecht). In deze projecten is voor de implementatie gebruikgemaakt van de gecombineerde top-down- en bottom-upmethode (Grol & Wensing 2006). Dat wil zeggen dat er een duidelijke en actieve ondersteuning is vanuit de top van de organisatie en dat men tegelijkertijd voor de keuze van de meetinstrumenten, de procedures en het gebruik van de instrumenten een beroep doet op de kennis van de professionals die ermee gaan werken.

De implementatie in de kortdurende ambulante psychiatrie bleek bij al deze projecten het weerbarstigst, vanwege de afstemming met de zorglogistiek, met aanmelding, intake, start en

beëindiging van behandeling (Buwalda 2011). In alle projecten is steeds kleinschalig gestart met een beperkt aantal teams om de procedure eerst te optimaliseren, alvorens deze 'uit te rollen' over de hele organisatie. Nuttig is het aanstellen van aandachtfunctionarissen die dicht bij de werkvloer werkzaam zijn en tot taak hebben te bewaken of de implementatie loopt zoals beoogt en problemen te signaleren.

In het ROM-project van GGZ Nederland worden keuzes gemaakt over de meetinstrumenten en de te volgen procedure, ten behoeve van de benchmark. De kans bestaat dat instellingen wachten met implementatie tot het ROM GGZ-project klaar is (Delespaul 2010). Dat kan ook betekenen dat professionals geen ruimte meer krijgen voor eigen invullingen van ROM. Het is voor de implementatie juist dan van belang dat instellingen ervoor zorgen dat professionals betrokken blijven en meedenken over een ROM-methodiek. Bij voorkeur zo dat zowel benchmarking mogelijk is als ook eigen variaties die recht doen aan de specifieke omstandigheden waaronder wordt gewerkt.

Bij de implementatie moet er aandacht zijn voor trainingsmogelijkheden (Nugter 2008; Trauer & Coombs 2010). Daarbij gaat het om (herhaalde) trainingen in betrouwbaar scoren van beoordelingsinstrumenten en om trainingen in het gebruik van de meetinstrumenten. Vaak veronderstelt men dat behandelaars wel weten wat ze moeten doen met meetinstrumenten. Dat blijkt vaak niet zo te zijn. Het gaat om gecompliceerde data die meerdere interpretaties mogelijk maken. Hafkenscheid demonstreert elders in dit nummer hoe uitkomsten van meetinstrumenten op uiteenlopende wijze kunnen worden geïnterpreteerd.

Voor veel professionals zal het wettelijk wettelijk zijn om samen met patiënten de voortgang mede te bespreken aan de hand van cijfers. Sommige professionals zien de bespreking met de patiënt als een verrijking, anderen als een bedreiging (De Jong & Nugter 2011). Voor die verschillen moet er aandacht zijn binnen organisaties.

Wat voorts helpt bij de implementatie is het zo snel mogelijk ‘teruggeven’ van de geaggregeerde gegevens aan de professionals. Presentaties over de eerste uitkomsten werken stimulerend (Nugter 2008; Nugter & Teer 2011).

Ten slotte is het van belang dat gegevens gemakkelijk opgenomen kunnen worden in elektronische patiëntendossiers. Niets is zo onpraktisch als het ene scherm te moeten openen om in het dossier te werken en voor de uitkomsten op de meetinstrumenten een ander scherm. Hoewel koppelingen tussen de systemen gemakkelijk worden beloofd, blijken die in de praktijk minder makkelijk te realiseren.

CONCLUSIES

ROM kan de ggz in potentie veel bieden. Dat heeft vooral te maken met de structuur ervan: door steeds op dezelfde manier een patiënt te evalueren, is beter vast te stellen of deze vooruitgaat of niet. Men kan dan sneller ingrijpen als dit tegen de verwachting in niet zo is. Voorwaarde is wel dat de scores op de meetinstrumenten ook echt gebruikt worden in een cyclus van plannen, uitvoeren, toetsen en opnieuw plannen. Voor professionals is dit nieuw. Ondersteuning en ruimte om daarmee te leren omgaan zijn essentieel.

Omdat bij grote groepen patiënten vergelijkbare meetinstrumenten worden gebruikt om het beloop te volgen, is het mogelijk om gegevens te aggregeren. Dit biedt de ggz veel aanknopingspunten voor verdere kwaliteitsverbetering. Ten eerste, omdat het primaire proces transparanter wordt. Ten tweede, omdat de mogelijkheid bestaat afdelingen en instellingen onderling met elkaar te vergelijken. Zo kan men van elkaar leren. Ten slotte bieden de geaggregeerde gegevens mogelijkheden voor nieuw onderzoek. Uiteindelijk zal dit weer leiden tot betere zorg.

LITERATUUR

- Buwalda VJA. Implementatie van meetinstrumenten in de kortdurende poliklinische behandeling. In: Buwalda VJA, Nugter MA, Swinkels JA, Mulder CL, red. *Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten*. Utrecht: De Tijdstroom; 2011. p. 103-20.
- Buwalda VJA, Nugter MA, Swinkels JA, Mulder CL, editors. *Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten*. Utrecht: De Tijdstroom; 2011.
- Beurs E de, Zitman F. Routine Outcome Monitoring. Het meten van therapie-effect in de klinische praktijk met webbased software. *Maandblad Geestelijke volksgezondheid* 2007; 62: 13-28.
- Jong K de, Nugter A. De kunst van het effectief feedback geven: waarom weten niet altijd tot verbeteren leidt. In: Buwalda VJA, Nugter MA, Swinkels JA, Mulder CL, red. *Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten*. Utrecht: De Tijdstroom; 2011. p. 51-64.
- De Jong K, Nugter MA, Polak MG, Wagenborg JEA, Spinhoven P, Heiser WJ. The Outcome Questionnaire (OQ-45) in a Dutch Population: A Cross-Cultural Validation. *Clin Psychol Psychoth* 2007; 14: 288-301.
- Delespaul AEG. De genormaliseerde T-score (2). Reactie op Edwin de Beurs, ‘De genormaliseerde T-score’. *Maandblad Geestelijke volksgezondheid* 2010; 65: 841-3.
- Deming WE. *Out of the crisis*. Cambridge, Massachusetts: Massachusetts Institute of Technology; 1982.
- Dickey B. Outcome measurement. From research to clinical practice. In: Ishak WW, Burt T, Sederer LI, red. *Outcome measurement in psychiatry, a critical review*. Washington: American Psychiatric Publishing; 2005. p. 15-22.
- Donabedian A. Evaluating the Quality of Medical care. *Milbank Mem Fund Q* 1966; 44 (Suppl): 194-6.
- Ellwood PM. Outcomes management. A technology of patient experience. 1988; 318: 1549-56.
- Grol RTPM, Wensing MJP. *Implementatie. Effectieve verbetering van de patiëntenzorg*. Amsterdam: Reed Business; 2006.
- Groot AD de. *Methodologie. Grondslagen van onderzoek en denken in de gedragswetenschappen*. 's Gravenhage: Mouton & Co; 1961.
- Hafkenscheid, A. De Outcome rating scale (ORS) en de Session rating scale (SRS): enkele psychometrische kenmerken van de Nederlandse versies. *Tijdschr Psychoth* 2010a; 36: 394-403.
- Hafkenscheid, A. Rammelende ROM in de ggz: geen ROM zonder Routine Process Monitoring (RPM). *GZ-Psychologie* 2010b; 2: 12-7.

- Hansen NB, Lambert MJ, Forman EV. The psychotherapy dose-response effect and its implications for treatment delivery services. *Clin Psychol Sci Prac* 2002; 9: 329-43.
- Hermann RC. Linking outcome measurement with process measurement for quality improvement. In: Ishak WW, Burt T, Sederer LI, red. *Outcome measurement in psychiatry, a critical review*. Washington: American Psychiatric Publishing; 2005. p. 23-34.
- Hooghiemstra TFM. *Privacy & managed care. Achtergrondstudies en Verkenningen*. Den Haag: Registratiekamer; 1998.
- Kerlinger FN. *Foundations of behavioral research*. New York: Holt, Rinehart and Winston; 1986.
- Lambert, M. Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychoth Res* 2007; 17: 1-14.
- Lambert MJ, Morton JJ, Hatfield DR, Harmon C, Hamilton S, Shimokawa K, e.a. *Administration and scoring manual for the OQ-45.2 (Outcome Questionnaire)*. Wilmington: American Professional Credentialing Services; 2004.
- Lambert MJ, Ogles BM. The efficacy and effectiveness of psychotherapy. In: Lambert ML, red. *Bergin and Garfield's handbook of psychotherapy and behaviour change*. New York: Wiley; 2004. p. 139-93.
- Lambert MJ, Whipple JL, Smart DW, Vermeersch DA, Nielsen SL, Hawkins, EJ. The effects of providing therapists with feedback on patient progress during psychotherapy: Are outcomes enhanced? *Psychoth Res* 2001; 11: 49-68.
- Miller SD, Duncan BL, Brown, J, Sparks J, Claud J. The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *J Brief Therapy* 2003; 2: 91-100.
- Miller SD, Duncan BL, Sorrel R, Brown GS. The Partners for Change Outcome Management System. *J Clin Psychol* 2005; 186: 199-208.
- Mulder CL, Staring ABP, Loos J, Buwalda VJA, Kuijpers D, Sytema S, Wierdsma AI. De Health of the Nation Outcome Scales (HoNOS) als instrument voor 'routine outcome assessment'. *Tijdschr Psychiatr* 2004; 46: 273-84.
- Nugter, A. Praktijkgericht onderzoek in de GGZ. Een literatuurverkenning. *Maandblad Geestelijke volksgezondheid* 2000; 55: 231-43.
- Nugter, A. De Routine Outcome Monitoring in FACT-wijktams. Implementatie en gebruiksmogelijkheden. In: Van Veldhuizen R, Bahler M, Polhuis D, Van Os J, red. *Handboek FACT*. Utrecht: De Tijdstroom; 2008. p. 255-72.
- Nugter MA, Van Bragt PM, Kumeling I. Uitkomsten van (deeltijd)klinische psychotherapie. Ervaren met een standaard evaluatieprocedure. *Maandblad Geestelijke volksgezondheid* 1998; 53: 861-74.
- Nugter MA, Buwalda VJA, Swinkels JA, Mulder CL. Kanttekeningen bij (extern) gebruik van meetinstrumenten. In: Buwalda VJA, Nugter MA, Swinkels JA, Mulder CL, red. *Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten*. Utrecht: De Tijdstroom; 2011. p. 189-92.
- Nugter MA, Teer W. Implementatie van meetinstrumenten in de langdurende psychiatrie. In: Buwalda VJA, Nugter MA, Swinkels JA, Mulder CL, red. *Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten*. Utrecht: De Tijdstroom; 2011. p. 131-46.
- Slade, M. Editorial: Outcome assessment in mental health services. *Psychol Med* 2002; 32: 1139-43.
- Sytema S, Nugter MA, Mulder CL. Routine outcome monitoring en onderzoek. In: Buwalda VJA, Nugter MA, Swinkels JA, Mulder CL, red. *Praktijkboek ROM in de ggz; een leidraad voor gebruik en implementatie van meetinstrumenten*. Utrecht: De Tijdstroom; 2011. p. 173-88.
- Trauer T, Coombs T. Routine outcome measurement: perspectives on skills and training. In: Trauer T, red. *Outcome measurement in mental health. Theory and practice*. Cambridge: University Press; 2010. p. 219-29.
- Walburg JA, Brinkmann J. *Kwaliteit is geen toeval: de praktijk van de kwaliteitszorg in de ggz*. Utrecht: GGZ Nederland; 2001.
- Wiger DE, Solberg KB. *Tracking mental health outcomes: a therapist's guide to measuring client progress, analyzing data, and improving your practice*. New York: Wiley; 2001.
- Wing JK, Beevor AS, Curtis RH, Park SB, Hadden S, Burns A. Health of the Nation Outcome Scales (HoNOS). Research and development. 1998; 172: 11-8.

AUTEURS

ANNET NUGTER is hoofd afdeling Onderzoek en Monitoring, GGZ Noord-Holland-Noord, Heerhugowaard.
 VICTOR BUWALDA is afdelingspsychiater, Altrecht, Utrecht.
 Correspondentieadres: dr. Annet Nugter, afdeling Onderzoek en Monitoring, GGZ NHN, Postbus 18, 1850 BA Heiloo.
 E-mail: a.nugter@ggz-nhn.nl

Geen strijdige belangen meegedeeld.

Het artikel werd voor publicatie geaccepteerd op 1-8-2011.

SUMMARY

Background and possible use of ROM in mental health care – M.A. Nugter, V.J.A. Buwalda –

BACKGROUND Routine outcome measurement or monitoring (ROM) has attracted worldwide attention largely as a result of developments in the field of quality care and scientific research.

AIM To provide insight into the significance of ROM for Dutch mental health care.

METHOD We consulted scientific literature on the following subjects: research and its applicability for clinical practice, quality improvement, ROM instruments and their implementation.

RESULTS ROM means that the outcomes of treatment and care are measured routinely as part of the plan-do-check-act cycle aimed at quality improvement. Measurement instruments are selected on clinical, scientific and practical grounds. ROM can be used in four main ways: for the treatment of individual patients, for the testing of policy, for benchmarking and for research. With regards to the implementation of ROM, a combined top-down, bottom-up approach is preferable to an entirely top-down or bottom-up approach.

CONCLUSION ROM can help to improve the quality of mental health care because it involves the standardized use of outcome measures and comprehensive databases.

[TIJDSCHRIFT VOOR PSYCHIATRIE 54(2012)2, 111-120]

KEY WORDS benchmarking, implementation, outcome research, PDCA-cycle, routine outcome monitoring, treatment planning