

ZICHTBARE ZORG GEESTELIJKE GEZONDHEIDSZORG

Managementsamenvatting

*Onderzoek naar de validiteit, betrouwbaarheid,
toepassing en gebruik van de
basisset prestatie-indicatoren GGZ*

Inleiding

In deze uitgebreide managementsamenvatting worden het onderzoeksproces en de resultaten besproken van het evaluatieonderzoek naar de prestatie-indicatoren voor de geestelijke gezondheidszorg en verslavingszorg. MediQuest voerde het onderzoek in 2008 uit in opdracht van de stuurgroep Zichtbare Zorg.. Hoofdvragen in het onderzoek waren:

1. Hoe valide zijn de prestatie-indicatoren in de basisset GGZ?
2. Hoe volledig zijn de door de zorgaanbieders in het Jaardocument Maatschappelijke Verantwoording (JMV) aangeleverde data?
3. Hoe betrouwbaar zijn de door de zorgaanbieders in het JMV aangeleverde data?
4. Hoe valide zijn de door de zorgaanbieders in het JMV aangeleverde data?
5. In welke mate is de informatie bruikbaar voor alle partijen?

De evaluatiestudie richt zich op 26 van de 28 indicatoren zoals opgenomen in de basisset 2007 – 2008. De indicatoren voor de forensische psychiatrie en de indicatoren 1.9 (Hanteren behandelrichtlijnen) en 2.5 (Suïcide) waren geen onderwerp van deze studie.

Gedurende het onderzoek leverde MediQuest diverse (tussen)rapportages en producten:

- Tussenrapportage 1 (TR1) voor het onderdeel validiteit: deze rapportage was ook input voor expertmeetings in september 2008. TR1 bevat een literatuurstudie waarin ook opmerkingen uit het veld zijn meegenomen (25 augustus 2008)
- Tussenrapportage 2 (TR2) voor het onderdeel validiteit: verslag van de bijeenkomsten met experts over effectiviteit, veiligheid en cliëntgerichtheid van de zorg, gehouden op 4, 8, en 11 september 2008 (29 september 2008);
- Eindrapportage deel 1: validiteit (13 oktober 2008)
- Geschoond databestand prestatie-indicatoren GGZ, bijbehorend tabellenboek en addendum (19 september 2008)
- Samenvattend verslag Focusgroep 4 november 2008: bruikbaarheid basisset prestatie-indicatoren GGZ en verslavingszorg voor cliëntenorganisaties
- Eindrapportage deel 2: volledigheid, betrouwbaarheid en bruikbaarheid

Na een korte schets van de achtergrond van dit evaluatieonderzoek worden per onderzoeksvraag beknopt de methode(n) en belangrijkste onderzoeksresultaten beschreven met tot slot een overall conclusie.

Achtergrond

Om de kwaliteit van de geestelijke gezondheidszorg transparant te maken, hebben negen partijen¹ de handen ineen geslagen en zich verenigd in de stuurgroep Zichtbare Zorg GGZ. Deze stuurgroep ontwikkelt, implementeert en borgt een basisset prestatie-indicatoren die de kwaliteit van de zorg inzichtelijk maakt. In 2007 hebben GGZ-instellingen op vrijwillige basis informatie aangeleverd voor de basisset prestatie-indicatoren. Vanaf 2008 zijn alle GGZ-aanbieders met een WTZi-erkenning verplicht deze informatie aan te leveren. Aanlevering vindt plaats in het JMV met behulp van de invoertool DigiMV.

De stuurgroep evalueert de prestatie-indicatoren periodiek en stelt deze waar nodig bij. In 2007 is een evaluatiestudie uitgevoerd² waarin de kwaliteit en de kwantiteit van de aanlevering van de

¹ Landelijk Platform GGZ, Inspectie voor de Gezondheidszorg, GGZ Nederland, V&VN, NVvP, NIP, NVVP, Ministerie van VWS, Zorgverzekeraars Nederland.

² Ham, van, M. A. J. A. E., Luijsterburg, J. C. & Brouwers, E. P. M. (2007). Evaluatie basisset prestatie-indicatoren voor de geestelijke gezondheidszorg en verslavingszorg, verslagjaar 2006, Tranzo, Universiteit van Tilburg.

gegevens van de prestatie-indicatoren is onderzocht. Medio juni 2008 heeft de stuurgroep aan onderzoeksbureau MediQuest opdracht gegeven een uitgebreid evaluatieonderzoek uit te voeren naar de validiteit, betrouwbaarheid, toepassing en het gebruik van de basisset prestatie-indicatoren GGZ.

Validiteit van de prestatie-indicatoren

Methodes

Om antwoord te kunnen geven op hoofdvraag 1 'Hoe valide zijn de prestatie-indicatoren in de basisset GGZ?' is een literatuurstudie verricht naar de relatie tussen de indicatoronderwerpen in de basisset en de kwaliteit van zorg. Ook is met behulp van experts in drie focusgroepen de indrukvaliditeit ('face validity') en de inhoudsvaliditeit ('content validity') nader onderzocht.

In totaal zijn zeven experts expliciet benaderd voor het aanleveren van relevante literatuur. Daarnaast zijn 33 experts benaderd voor deelname aan de focusgroepen en voor het aanleveren van referenties. Voorafgaand aan het verzamelen van relevante literatuur bij experts zijn in Pubmed en het Tijdschrift voor Psychiatrie aan de hand van vooraf vastgestelde zoekcriteria relevante referenties gezocht. De concept referentielijst is verzonden aan veertig experts met een verzoek tot aanvulling. In totaal zijn 70 referenties bestudeerd. Voor elke relevante referentie is een korte samenvatting gemaakt, die in Tussenrapportage 1 is opgenomen. Daarnaast zijn in deze rapportage de opvallende punten uit de toelichtingen van de instellingen bij de door hun aangeleverde data meegenomen. Tussenrapportage 1 vormde de input voor de drie focusgroepen rondom de thema's effectiviteit, veiligheid en cliëntgerichtheid met in totaal 17 experts. Een overzicht van de deelnemers en de besproken indicatoren per expertgroep is opgenomen in de bijlage.

Resultaten: relevantie indicatoronderwerpen

Voor vrijwel alle indicatoronderwerpen geldt dat in de literatuur onderbouwing is gevonden voor de relatie tussen het onderwerp en de kwaliteit van zorg. Een uitzondering hierop is indicator 2.3 (Onveiligheid door het ontbreken van informatie). Hiervoor is in dit onderzoek geen onderbouwende literatuur gevonden. In de expertgroep is dit indicatoronderwerp wel als een belangrijk kwaliteitsaspect van zorg benoemd. Een tweede uitzondering is prestatie-indicator 1.6 (Percentage heropname versus gemiddelde opnameduur). In de literatuur is slechts beperkt bewijs gevonden voor de relatie tussen heropname en de geleverde kwaliteit van zorg. Aanvullend onderzoek is vereist om deze relatie al dan niet aan te tonen.

Resultaten: indrukvaliditeit prestatie-indicatoren

In de focusgroepen zijn de onderbouwing, de vraagstelling en de exclusiecriteria van elke indicator kritisch geëvalueerd. In bijlage 3 van Eindrapportage 1 zijn voor elke afzonderlijke prestatie-indicator de belangrijkste bevindingen beschreven. De experts hebben na bespreking van elke indicator een oordeel gegeven over de indrukvaliditeit van de prestatie-indicator zoals deze nu is opgenomen in de basisset. De uitkomsten hiervan zijn weergegeven in kader 1³. Hieruit blijkt dat 8 van de 26 indicatoren volgens de experts geen of slechts geringe aanpassing behoeven om valide resultaten te meten (score 4 of 3). Elf indicatoren vereisen een meer uitgebreide aanpassing aangezien een deel van de experts van mening is dat de indicatoren zoals ze nu zijn geformuleerd slechts in beperkte mate iets zeggen over de door instellingen geleverde kwaliteit van zorg. De benodigde aanpassingen liggen in een aanscherping van de onderbouwing van de indicator, van de vraagstelling en/of wijziging van de exclusiecriteria. De indicatoren 1.1 (Bereik preventieactiviteiten), 1.3 (Verandering in het dagelijks functioneren van de cliënt), 1.5 (Rehabilitatie), 2.3 (Onveiligheid door het ontbreken van informatie) en 3.2 (Toegang tot zorg) worden door de experts, zoals ze in de huidige set zijn geoperationaliseerd, als niet valide beschouwd. De argumenten hiervoor zijn beschreven in bijlage 3 van Eindrapportage 1.

³ In Tussenrapportage 2 'Verslag focusgroepen met experts 4,8 en 11 september 2008', is de achterliggende discussie bij de oordelen weergegeven

Kader 1

HEEFT U OP BASIS VAN UW EXPERTISE (EN TR1) DE INDRUK DAT DE INDICATOR IETS ZEGT OVER DE DOOR INSTELLINGEN GELEVERDE KWALITEIT ('FACE VALIDITY')?

PI:	Oordeel
4	Ja / zeer eens
3	In meerdere mate / eens
2	In mindere mate / oneens
1	Nee / zeer oneens

EFFECTIVITEIT										
1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.1 0	1.1 1	

PI:	Oordeel
4	Ja / zeer eens
3	In meerdere mate / eens
2	In mindere mate / oneens
1	Nee / zeer oneens

VEILIGHEID				
2.1	2.2	2.3	2.4	2.6

PI:	Oordeel
4	Ja / zeer eens
3	In meerdere mate / eens
2	In mindere mate / oneens
1	Nee / zeer oneens

CLIENTGERICHTHEID										
3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.1 0	3.1 1

Toelichting:

Indien minimaal twee experts voor een bepaalde antwoordcategorie hebben gekozen is de bijbehorende cel ingekleurd. Uitzondering hierop is PI 3.6: hierbij zijn twee experts het oneens i.t.t. de overige experts, echter zij scoren in verschillende antwoordcategorieën (1 oneens, 1 zeer oneens). Om deze reden zijn de antwoordopties licht gearceerd.

Resultaten: algemene bevindingen over de validiteit van de basisset

Op basis van het literatuuronderzoek, de analyse van de toelichtingen door de instellingen op de aangeleverde indicatoren en de uitkomsten van focusgroepen met experts zijn er algemene aandachtspunten te benoemen die relevant zijn voor de validiteit van de huidige basisset.

Perspectief

- In de huidige basisset is vooral gekeken naar de haalbaarheid voor de instellingen om data voor een prestatie-indicator te kunnen opleveren. Dit gaat logischerwijs ten koste van de zeggingskracht. De set wordt stapsgewijs doorontwikkeld. De huidige set moet dan ook vanuit dit perspectief gezien worden. De basisset draagt bij aan het kritischer naar de zorg kijken, en aan bewustwording en kwaliteitsverbetering.
- De titel van het hoofdonderwerp effectiviteit in de basisset is misleidend. Deze suggereert dat de uitkomsten daadwerkelijk iets zeggen over de effectiviteit van zorg. Hiervoor is het nog te vroeg. Door het gebruik van verschillende meetinstrumenten en grote verschillen in de frequentie van meten, leveren de uitkomst prestatie-indicatoren 1.2 t/m 1.4 op dit moment nog onvoldoende valide en vergelijkbare resultaten op. De procesindicatoren behorend bij het hoofdonderwerp effectiviteit zeggen (na aanscherping) wel iets over de kwaliteit van zorg.

- Cliëntgerichtheid hoeft niet alleen gemeten te worden met vragen aan de cliënt. Vice versa geldt ook dat de andere hoofdonderwerpen, effectiviteit en veiligheid, behalve aan de hulpverlener, ook gevraagd (kunnen) worden aan de cliënt. Voor de indicatoren van het hoofdonderwerp cliëntgerichtheid wordt veelal uit slechts één bron geput: de schriftelijke cliëntervaringmeting (met behulp van de Thermometer cliëntwaardering of CQ-index).
- Kwaliteit kan niet alleen afgemeten worden aan de cliënten die je hebt, maar ook aan de mensen die je niet in zorg hebt (latente zorgbehoefte). Aandachtspunt is de allochtone (potentiële) klantgroep. Deze klantgroep wordt onvoldoende zichtbaar gemaakt met de 'prestatie-indicatoren' in de basisset, noch door de vragenlijsten van de CQ-index.

Onderbouwing

- Voor veel indicatoren geldt dat de onderbouwing van de keuze voor specifiek deze prestatie-indicator explicieter in de basisset opgenomen moet worden. Temeer wanneer de prestatie-indicator een smaller terrein bestrijkt dan de indicatornaam doet vermoeden.
- In de onderbouwing van prestatie-indicatoren 1.2 t/m 1.4 ontbreekt de vermelding dat op dit moment nog voorzichtigheid is geboden bij de interpretatie van de uitkomstindicatoren.
- Bij de prestatie-indicatoren 1.2, 1.4, 1.7, 1.11 en 2.4 worden in de onderbouwing uitspraken gedaan over onderzoeksresultaten zonder referenties te benoemen.
- Mogelijk geldt voor sommige indicatoren dat de meerwaarde van uitsplitsing in diagnosegroepen beperkt is (zoals bij indicator 1.4). Een onderbouwing voor de gevraagde uitsplitsing ontbreekt.

Definities

- In de inleiding van de basisset ontbreken de definities voor uitkomst- proces- en structuurindicatoren. Er is met name nog onduidelijkheid over wat wordt verstaan onder een uitkomstindicator. De prestatie-indicatoren 1.1 en 1.10 zijn mogelijk procesindicatoren in plaats van uitkomstindicatoren, alnaargelang de gehanteerde definitie.
- In de inleiding van de basisset ontbreken de definities voor effectiviteit, veiligheid, cliëntgerichtheid en kwaliteit van zorg.
- In de basisset ontbreken soms essentiële definities voor een goede begripsvorming van de prestatie-indicator. Deze definities zijn in sommige gevallen wel expliciet opgenomen in de werkinstructie. Een goed voorbeeld hiervan is prestatie-indicator 1.4 (Verandering in de ervaren kwaliteit van leven van de cliënt), waarvoor in de werkinstructie een definitie wordt gegeven van langdurige zorg en wordt gedefinieerd wanneer er sprake is van verbetering en stabilisatie. Deze definitie ontbreekt in de basisset.

Validiteit van de data

Methode

Om antwoord te kunnen geven op hoofdvraag 4: 'Hoe valide zijn de door de zorgaanbieders in het JMV aangeleverde data?' zijn de toelichtingen op de data die door de GGZ-instellingen zijn gegeven bij de twee aanleveringen in DigiMV (1 juni en 1 september 2008) geïnventariseerd en zijn diverse statische analyses op de data uitgevoerd om inzicht te krijgen in:

- de mate van spreiding van de aangeleverde gegevens op de individuele indicatoren
- hoeverre de huidige indicatorenset de potentie heeft om eventuele verschillen tussen verschillende typen zorginstellingen aan het licht te brengen
- de onderlinge samenhang tussen individuele indicatoren, zowel binnen als tussen de diverse hoofdonderwerpen

Resultaten: spreiding van de data

In Eindrapportage 1, bijlage 4 is per (sub)indicator de spreiding van de geschoonde data grafisch weergegeven (in boxplots). Over het algemeen zit er een behoorlijke spreiding in de aangeleverde data. Dat doet vermoeden dat de prestatie-indicatoren verschillen tussen individuele instellingen laten zien. Ten aanzien van de spreiding is er echter een aantal zaken waarmee bij de interpretatie van de data rekening moet worden gehouden. Hieronder worden de belangrijkste genoemd:

- Voor de cliëntenraadpleging worden verschillende meetinstrumenten gebruikt. Het merendeel van de instellingen heeft op basis van de Cliëntthermometer en/of een eigen meetinstrument gerapporteerd (slechts 9 instellingen hebben op basis van hun deelname aan de pilot CQ-index gerapporteerd). Hoewel de gemiddelde uitkomsten veelal vergelijkbaar zijn, laten de uitkomsten in geval van een eigen meetinstrument consistent een grotere spreiding zien dan die van de Cliëntthermometer. Dit is te wijten aan het feit dat als er bijvoorbeeld 60 instellingen zijn die een eigen meetinstrument hebben gehanteerd, dit in theorie 60 verschillende instrumenten kunnen zijn met elk hun eigen specifieke vraagstellingen en scoringsmethodieken. De validiteit van een aldus gemeten indicator komt door het ontbreken van uniformiteit in het meetproces in het geding.
- Onjuiste interpretatie van de gevraagde gegevens bij indicatoren gebaseerd op cliëntenraadpleging. De 9 instellingen die de indicatoren uit het onderdeel cliëntgerichtheid gebaseerd hebben op de resultaten uit de pilot met de CQ-index blijken nog onvoldoende bekend te zijn met de terminologie die wordt gehanteerd voor de uitkomstmaten van de CQ-index. Zo blijkt uit de data dat voor de indicatoren 1.2a, 3.4, 3.5, 3.7, 3.10 en 3.11 i.p.v. schaal-totaalscores andere scores zijn gegeven voor de teller, bijvoorbeeld gemiddelden. Daarnaast blijken sommige instellingen, die meten met de Cliëntthermometer of een eigen meetinstrument, de uitkomstmaten die gebaseerd zijn op meer dan één vraag (indicator 3.5, 3.10 en 3.11) anders te interpreteren dan bedoeld. Dat wil zeggen, in plaats van het aantal cliënten aan te leveren dat beide vragen positief heeft beantwoord, wordt dan de volgende optelling gemaakt: [aantal cliënten dat positief heeft geantwoord op vraag 1] + [aantal cliënten dat positief heeft geantwoord op vraag 2].
- Afwijkende definities. Uit de toelichtingen die instellingen geven blijkt dat sommige instellingen afwijkende definities hanteren voor indicator 1.6a (Percentage heropnames voor dezelfde diagnose/reden binnen drie maanden) en indicator 1.8 (Somatische screening). Er zijn instellingen die i.p.v. 3 maanden een termijn van 6 maanden hanteren of bijvoorbeeld juist alle heropnames meenemen ongeacht de tijd die is verstreken sinds de vorige opnameperiode. De definitie 'somatische screening door een bevoegd arts' en de termijn 'binnen 1 week na opname' wordt mogelijk niet door alle instellingen even consistent gevolgd.
- Gegevens gebaseerd op schattingen i.p.v. administratieve gegevens. Met name voor indicator 1.1 (bereik preventieactiviteiten) en indicator 1.10 (bereik zorgwekkende zorgmijders) geldt dat de gevraagde gegevens voor de noemer niet uit de zorgadministratie is te halen en het voor instellingen lastig is een reële schatting te geven. Dit gegeven wordt weerspiegeld in de aanzienlijke spreiding van de data en de aanwezigheid van een aantal extreme waarden (meer dan 3 standaarddeviaties verwijderd van het gemiddelde).

Waar nodig en mogelijk zijn data op basis van interpretatie van door instellingen gegeven toelichtingen geschoond. Vaak is echter geen toelichting gegeven, waardoor eventuele onjuistheden niet per definitie af te leiden zijn uit de data.

Resultaten verschillen tussen verschillende typen zorgaanbieders

Voor het identificeren van potentiële verschillen tussen verschillende typen zorgaanbieders is er een onderscheid gemaakt tussen zorgaanbieders die alleen ambulante behandeling/begeleiding bieden (ambulant), zorgaanbieders die uitsluitend begeleid of beschermd wonen bieden (RIBW), zorgaanbieders die uitsluitend klinische behandeling/begeleiding bieden (klinisch) en een restcategorie die bestaat uit de zorgaanbieders die een combinatie van twee of meer van voornoemde zorgsoorten in hun pakket hebben (geïntegreerd). In totaal was er bij 29 uitkomstmaten sprake van een significant effect van instellingstype. Dit impliceert dat de gemiddelden van de betreffende uitkomstmaten aantoonbaar verschillen tussen de verschillende typen instellingen. Voor een overzicht van alle (sub)indicatoren waarbij sprake is van een significant effect van instellingstype wordt verwezen naar bijlage 5 van Eindrapportage deel 1. Voorbeelden van een grote invloed (hoogsignificant) van instellingstype op de uitkomstmaten zijn:

- Het percentage cliënten binnen de diagnosegroep stemmingsstoornissen, waarbij het dagelijks functioneren bij voor- en nameting is verbeterd, is significant hoger binnen de categorie ambulant dan binnen de categorie geïntegreerd (indicator 1.3)
- De gemiddelde wachttijd voor ambulante zorg is significant korter voor instellingen die uitsluitend ambulante behandeling/begeleiding bieden dan voor instellingen die een combinatie van verschillende zorgsoorten leveren (indicator 3.1a)
- Het percentage cliënten dat keuzevrijheid ervaart ten aanzien van kiezen van de eigen behandelaar en kiezen uit geboden zorgaanbod is significant hoger voor instellingen uit de categorie ambulant dan voor instellingen uit de categorie geïntegreerd (indicator 3.4)

Resultaten: correlatie binnen en tussen hoofdonderwerpen

Voor de verschillende combinaties van (sub)indicatoren (met minimaal 30 geldige waarnemingen) is de correlatiecoëfficiënt (Pearson R) berekend. In totaal zijn er maar liefst 76 combinaties van (sub)indicatoren geïdentificeerd waarbij sprake was van een significante ($p < 0.05$) en sterke ($R > 0.7$) correlatie. Enige nuancering is hier echter op z'n plaats; het merendeel van deze combinaties betreft combinaties van dezelfde (sub)indicator, maar dan voor verschillende diagnosegroepen. In bijlage 6 van Eindrapportage deel 1 wordt een overzicht gegeven van alle combinaties waar sprake is van een significante correlatie die betrekking heeft op verschillende indicatoren en niet slechts op verschillende diagnosegroepen binnen dezelfde indicator. Een voorbeeld hiervan is de positieve samenhang tussen indicator 1.2a (percentage cliënten dat aangeeft dat behandeling heeft geleid tot verbetering/stabilisatie van de klachten, gemeten met Cliëntthermometer) en 3.5 (percentage cliënten dat aangeeft dat behandeling/begeleiding voldoet aan eigen zorgwensen, gemeten met Cliëntthermometer).

De samenhang tussen fundamenteel verschillende indicatoren is voorsnog wat beperkt. Dit wil echter niet per definitie zeggen dat de huidige set indicatoren onvoldoende samenhangend is; het is zeer goed mogelijk dat de beperkte aanlevering van bepaalde indicatoren, met name uit het hoofdonderwerp veiligheid, debet is aan het ontbreken van meer significante correlaties. Een hopelijk meer volledige en betrouwbare aanlevering over verslagjaar 2008 zal hier naar verwachting uitsluitel over kunnen geven.

Volledigheid

Methode

Om antwoord te geven op hoofdvraag 2 'Hoe volledig zijn de door de zorgaanbieders in het JMV aangeleverde data?' is een driedelige kwantitatieve analyse verricht naar:

- de volledigheid op instellingsniveau
- de volledigheid op indicatorniveau
- de volledigheid ten aanzien van het gesplitst per diagnosegroep aanleveren van indicatoren waarvoor dat van toepassing is

Hiervoor is per instelling, per uitkomstmaat (o.b.v. zorgtype en diagnosegroepen) bepaald of de betreffende uitkomstmaat aangeleverd had moeten/kunnen worden. Vervolgens is per instelling, per uitkomstmaat bepaald of de betreffende uitkomstmaat ook daadwerkelijk in het geschoonde bestand beschikbaar is. Daarnaast is het onderwerp volledigheid op een meer kwalitatieve manier in kaart gebracht. Om inzicht te verschaffen in de achterliggende redenen voor het al dan niet aanleveren van bepaalde indicatorengegevens zijn de toelichtingen van instellingen in de dataset bestudeerd en zijn met kwaliteitsfunctionarissen van vijf GGZ-instellingen interviews gehouden.

Resultaten

In het databestand van 2008 zijn in totaal 168 instellingen opgenomen. Bij de analyse bleek dat voor 36 van de 168 instellingen gold dat deze geen enkele indicator over verslagjaar 2007 hebben aangeleverd. Bij de analyses voor de volledigheid zijn deze instellingen buiten beschouwing gelaten.

Gemiddeld genomen moeten de geïntegreerde GGZ-instellingen een groter aantal indicatoren van de basisset aanleveren dan de meeste andere instellingstypen. De geïntegreerde GGZ-instellingen leveren gemiddeld genomen ook het meest volledig aan, zoals blijkt uit tabel 1. Daarin is een aantal statistische grootheden m.b.t. het aanleverpercentage per instellingstype is opgenomen.

Tabel 1. Overzicht van het aantal daadwerkelijk aangeleverde uitkomstmaten als percentage van het aantal aan te leveren uitkomstmaten, als functie van instellingstype

Instellingstype	Gemiddelde (%)	Std. Deviatie (%)	Minimum (%)	Maximum (%)	Mediaan (%)	N
Geïntegreerde GGZ instellingen	53,0	16,7	7,5	94,8	54,1	38
RIAGG	43,5	7,8	34,5	52,2	43,7	4
Overige instellingen voor ambulante GGZ	26,6	17,4	4,2	54,7	21,4	14
RIBW	20,7	13,2	2,6	42,4	16,3	13
Gespecialiseerde instellingen voor verslavingszorg	44,5	19,5	10,4	66,7	52,1	8
Gespecialiseerde instellingen voor kind en jeugd	36,3	18,5	11,1	57,3	38,3	9
PAAZ	24,7	20,0	2,0	57,0	22,3	6
Overig	26,7	22,9	2,5	73,3	15,2	40

De indicatoren voor veiligheid zijn het beste aangeleverd; gemiddeld 73.6% van de instellingen die de betreffende indicatoren had moeten/kunnen aanleveren, heeft deze ook daadwerkelijk aangeleverd. Het gemiddelde percentage voor cliëntgerichtheid bedraagt 63.8% en dat voor

effectiviteit 37.5%. Het relatief lage aanleverpercentage voor effectiviteit is hoofdzakelijk toe te schrijven aan het feit dat een groot deel van de daartoe behorende indicatoren per diagnosegroep diende te worden aangeleverd, wat voor een deel van de instellingen niet mogelijk bleek te zijn.

Redenen die meerdere instellingen noemden voor het onvolledig aanleveren worden hieronder besproken:

- Uitsplitsing per diagnosegroep is nog niet mogelijk. Sommige instellingen geven aan t.z.t. de uitsplitsing voor de gegevens over verslagjaar 2008 wel te kunnen maken.
- Over indicator 1.3 (Verandering in het dagelijks functioneren van de cliënt, gemeten met het meetinstrument de GAF) wordt regelmatig opgemerkt dat men moeite heeft met het correct toepassen van het GAF-instrument, of dat men twijfels heeft bij de betrouwbaarheid ervan, en dat gegevens voor deze indicator om die reden niet beschikbaar zijn of wel beschikbaar zijn maar niet worden aangeleverd.
- Ten aanzien van indicator 2.3 (Onveiligheid door ontbreken van informatie) geven instellingen te kennen geen onderscheid te (kunnen) maken tussen 'binnen kantooruren' en 'buiten kantooruren', en dat zij om die reden de betreffende indicator niet kunnen aanleveren.
- Bij indicator 3.1 (Wachttijd tot start behandeling) geven instellingen regelmatig aan bij de aan te leveren gegevens geen onderscheid te kunnen maken tussen ambulante en klinische, waardoor er óf totalen worden aangeleverd óf de gegevens voor beide categorieën niet worden aangeleverd.
- Bij indicatoren die gebaseerd zijn op specifieke vragen uit de CQ-index, maar dan gemeten met vergelijkbare vragen uit een eigen meetinstrument geven instellingen zo nu en dan aan dat de door hun gestelde vragen dusdanig afwijken van de CQ-index-vragen dat ze om die reden de betreffende indicator niet hebben aangeleverd.

Betrouwbaarheid

Methoden

Om antwoord te kunnen geven op hoofdvraag 3 'Hoe betrouwbaar zijn de door de zorgaanbieders in het JMV aangeleverde data?' zijn de toelichtingen die de instellingen gaven bij het aanleveren van de gegevens geïnventariseerd en gecategoriseerd. De inventarisatie van de toelichtingen vormde samen met een aantal kwantitatieve aspecten input voor het opstellen van te hanteren schoningsregels (voor data die de instellingen hebben aangeleverd). Voorbeelden van schoningsregels zijn het verwijderen van onwaarschijnlijke waarden (bijvoorbeeld percentages > 100%), uitbijters en gegevens die zijn gebaseerd op een schatting. Een overzicht van alle gehanteerde schoningsregels is weergegeven in bijlage 3 van Eindrapportage deel 2.

Resultaten

Bij de bevindingen over de validiteit van de data zijn een aantal situaties benoemd die invloed hebben op de validiteit en daarmee ook op de betrouwbaarheid van de data. Voorbeelden hiervan zijn het niet strikt hanteren van definities van de 'prestatie-indicatoren uit de basisset en een onjuiste interpretatie van de vraagstelling. Uit de toelichtingen van instellingen is ook gebleken dat met name bij indicator 1.7 (Drop out) sommige instellingen de in- en exclusiecriteria verschillend interpreteren. Bij het onderdeel volledigheid is de problematiek rondom het meetinstrument de GAF bij indicator 1.3 reeds besproken. Het spreekt voor zich dat een onbetrouwbaar meetinstrument ten koste gaat van de betrouwbaarheid van de indicator als geheel.

Bij de kwantitatieve analyse van de data zijn er in totaal 34 uitbijters, verspreid over 27 uitkomstmaten, geïdentificeerd en verwijderd (voor een overzicht van de uitkomstmaten waarin uitbijters zijn geïdentificeerd zie bijlage 5). Op een totaal van 3944 waarnemingen verdeeld over 152 uitkomstmaten komt dat neer op een percentage van slechts 0.9%. Het relatief geringe aantal uitbijters dat aan het licht is gekomen impliceert echter niet per definitie dat de data

zonder meer in hoge mate betrouwbaar zijn. Immers, de vaak variabele interpretatie van indicatoren en het niet strikt hanteren van de voorgeschreven in- en exclusiecriteria door de instellingen heeft er onherroepelijk toe geleid dat de spreiding van de data relatief groot is. Hierdoor wordt de kans op het identificeren van uitbijters aanzienlijk gereduceerd.

Bruikbaarheid

Method

Om antwoord te kunnen geven op hoofdvraag 5 'In welke mate is de informatie bruikbaar voor alle partijen?' zijn semigestructureerde interviews gehouden met vijf kwaliteitsfunctionarissen van GGZ-instellingen, vier eerste geneeskundigen, drie zorginkopers van verzekeringsmaatschappijen en twee IGZ-inspecteurs voor de GGZ. Daarnaast is een focusgroep bijeenkomst voor cliëntenvertegenwoordigers georganiseerd in samenwerking met het Landelijk Platform GGZ (LPGGz). In totaal hebben 23 vertegenwoordigers van landelijke, regionale en lokale cliëntenorganisaties in de GGZ en verslavingszorg aan de bijeenkomst deelgenomen.

Resultaten

Voorafgaand aan de bijeenkomst hebben de cliëntenvertegenwoordigers een individuele top 5 samengesteld van aspecten waarop een GGZ-instelling beoordeeld zou moeten worden. Uit de geaggregeerde top 5-lijsten blijkt dat meer dan de helft van de genoemde onderwerpen op dit moment ook onderdeel uitmaakt van de basisset, te weten: continuïteit van zorg, evaluatie behandeling, bejegening, informatieverstrekking, regie/zeggenschap cliënten, omgevingsveiligheid, drop out, rehabilitatie, wachttijden, behandelresultaat en kwaliteit van de somatische zorg. Wel zijn er verschillen in de wijze waarop de onderwerpen in prestatie-indicatoren zijn geoperationaliseerd en de door cliëntenorganisaties aangedragen onderwerpen. Zo omvat het onderwerp 'regie/zeggenschap cliënten' behalve aspecten ten aanzien van de keuzevrijheid ook aspecten ten aanzien van de positie en invloed van de cliëntenraad in de organisatie.

Daarnaast zijn er onderwerpen genoemd die niet voorkomen in de huidige basisset:

- inzet en samenwerking met ervaringsdeskundigen
- deskundigheid personeel
- contact met en ondersteuning aan omgeving cliënt
- perspectief van de allochtone cliënt

Cliëntenvertegenwoordigers hebben aangegeven de prestatie-indicatoren behalve als keuze ondersteunende informatie ook te willen gebruiken voor belangenbehartiging. De gegevens van de instellingen op de prestatie-indicatoren uit de basisset kunnen als informatiebron dienen in de contacten met zorgaanbieder en zorgverzekeraar. Wel ontbreekt het nog aan een toegankelijk, gebruiksvriendelijk overzicht van deze gegevens. Ook zijn veel cliëntenvertegenwoordigers nog niet op de hoogte van het bestaan van de basisset prestatie-indicatoren en de verplichting voor instellingen om hierover gegevens aan te leveren.

De inspectie wil de gegevens uit de basisset gebruiken voor het gefaseerd toezicht. Het doel om de inspecteurs van goede toereikende informatie te voorzien voor hun taakuitoefening, is naar de mening van de twee geïnterviewde inspecteurs nog niet gehaald. Op het moment van het onderzoek had het kenniscentrum van de IGZ de gegevens van de instellingen over het jaar 2007 nog niet verspreid onder de inspecteurs. In 2007 hebben de inspecteurs voor het eerst informatie gekregen over 78 organisaties over verslagjaar 2006.

De zorginkopers maken als enige al uitgebreid gebruik van de prestatie-indicatoren om a) zorginstellingen te prikkelen tot kostenbewustzijn en kwaliteitsverhoging en b) te differentiëren in de zorginkoop. De inkopers maken hiertoe gebruik van uitsluitend goed meetbare indicatoren, voornamelijk vooral op het gebied van cliëntgerichtheid.

Hoewel de geïnterviewde eerste geneeskundigen de onderwerpen in de basisset in meer of mindere mate belangrijk vinden, maken zij nog geen gebruik van de uitkomsten per indicator. Zij benadrukken dat er de laatste jaren enkele stuk voor stuk goede initiatieven zijn genomen, waaronder de introductie van de basisset prestatie-indicatoren, de HKZ-certificering en nu ook de Routine Outcome Monitoring. Men stelt zich hierbij wel de vraag of al deze initiatieven niet een te zware last vormen voor de instellingen. Er is behoefte aan coördinatie van en afstemming tussen de verschillende kwaliteitsinitiatieven.

Drie van de vijf geïnterviewde kwaliteitsfunctionarissen (deze kwaliteitsfunctionarissen werken bij andere GGZ-instellingen dan de geïnterviewde eerste geneeskundigen) geven aan dat prestatie-indicatoren intern worden gebruikt voor sturing. Dit komt bijvoorbeeld tot uiting door het opnemen van normen voor enkele 'prestatie-indicatoren in managementafspraken.

Zorgaanbieders (eerste geneeskundigen en kwaliteitsfunctionarissen) denken verschillend over de kwaliteit van de huidige basisset voor externe verantwoording. Het publiceren van de prestatie-indicatoren per instelling vergroot de vergelijkbaarheid van de instellingen. Het vergelijken van de instellingen met behulp van de prestatie-indicatoren moet wel met de nodige voorzichtigheid geschieden. Instellingen en hun cliënten verschillen immers onderling sterk, er zijn interpretatieverschillen over enkele prestatie-indicatoren en er is slechts beperkte externe controle is op de verzamelde data.

Tot slot

De komende jaren worden de prestatie-indicatoren doorontwikkeld op basis van evaluatiestudies en gebruikerservaringen. Daarbij is het niet de bedoeling om meer indicatoren vast te stellen, maar om de gekozen indicatoren aan te scherpen en te verbeteren. Een ander belangrijk uitgangspunt is dat er altijd een balans gezocht moet worden tussen zeggingskracht enerzijds en efficiëntie en haalbaarheid anderzijds. Dat kan betekenen dat omwille van de hanteerbaarheid en leverbaarheid van de indicatoren soms voor een indicator gekozen wordt die niet optimaal valide en betrouwbaar is. De uitkomsten van het evaluatieonderzoek naar de validiteit, betrouwbaarheid, toepassing en gebruik van de basisset prestatie-indicatoren GGZ (samengevat in 25 concrete conclusies en aanbevelingen in de eindrapportage) dienen in dit licht gezien te worden.

Uit het onderzoek blijkt dat de stuurgroep Zichtbare Zorg GGZ over het algemeen de juiste indicatoronderwerpen heeft geselecteerd; onderwerpen die op basis van de verrichte literatuurstudie en/of naar de mening van de experts en gebruikers belangrijk zijn voor het in kaart brengen van de kwaliteit van de geestelijke gezondheidszorg en verslavingszorg. Veel verbeterpunten liggen op het vlak van de operationalisatie van de indicatoren, zowel in de onderbouwing, de vraagstelling, als in de definiëring van de doelgroep.

Waren er in 2007 78 instellingen die, al dan niet volledig, gegevens voor de basisset hebben aangeleverd, in 2008 is dit aantal bijna verdubbeld. De verwachting is dat verdere doorontwikkeling van de basisset en van de wijze van dataverzameling binnen de instellingen zal resulteren in jaarlijks meer volledige, valide, betrouwbare en bruikbare gegevens.

BIJLAGE DEELNEMERS EXPERTGROEPEN

Facilitator: Drs. L. Roggeveen, directeur Evaluatiegroep Overheid en Bedrijf
Verslaglegging: Mw. drs. A.F. Vissers, onderzoeker MediQuest

Deelnemers Expertgroep 4 september 2008, Effectiviteit

Naam	Functie	Organisatie
Prof. dr. J. Dekker	Hoofd Onderzoek/ Hoogleraar Evidence Based behandeling	Jellinek Mentrum / Vu Amsterdam
Mw. drs. M.P.E. Janssen	Inspecteur	IGZ
Dr. C.G. Schoemaker	Senior onderzoeker	RIVM (Zorgbalans)
Mw. drs. M. Smolders	Junior onderzoeker/ promovendus	IQ healthcare
Dr. P.F.M. Verhaak	Programmaleider GGZ	Nivel
P. Lamers*	Inspecteur	IGZ

Prestatie-indicatoren 4/9: 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.8, 1.10, 1.11, 3.2, 3.7.

Deelnemers Expertgroep 8 september 2008, Veiligheid (/Effectiviteit)

Naam	Functie	Organisatie
Mw. dr. M.C. de Bruijne	Onderzoeker	EMGO/VUMC
Prof. dr. A.J.M. Janssen	Bestuurder/ Onderzoeker/ Hoogleraar Economie van de Gezondheidszorg	VVT / Mondriaan/ Universiteit van Tilburg (Tranzo)
W.A. Janssen Msc.	Onderzoeker	GGNET
Mw. drs. M.P.E. Janssen	Inspecteur	IGZ
Prof. dr. A.J.M. Loonen	Hoogleraar Farmacotherapie bij psychiatrische patiënten	RUG
Mw. drs. E. Poot	Adviseur	LEVV
Mw. drs. M. Smolders	Junior onderzoeker/ promovendus	IQ healthcare
Mw. dr. K. Taxis	Onderzoeker	RUG
P. Lamers*	Inspecteur	IGZ
Mw. J. van den Hoek*	Beleidsmedewerker	GGZ

Prestatie-indicatoren 8/9: 2.1, 2.2, 2.3, 2.4, 2.6, 1.5, 1.6, 3.2.

Deelnemers Expertgroep 11 september 2008, Cliëntgerichtheid

Naam	Functie	Organisatie
Drs. H.A.P. Beijers	Manager	Steunpunt GGZ
Mw. Prof. dr. D.M.J. Delnoij	Directeur	Centrum Klantervaring Zorg
Dr. H.P.M. van Haaster	Onderzoeker/ training	I.G.P.B.
Mw. drs. M.A. Janssen	Inspecteur	IGZ
Drs. S.H. Makkink	Onderzoeker	Kwadraad
P. Lamers*	Inspecteur	IGZ

Prestatie-indicatoren 11/9: 3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10, 3.11

* Dhr. P. Lamers en mw. J. van den Hoek hebben niet als expert maar als vertegenwoordiger van de projectgroep Transparantie GGZ deelgenomen aan de expertgroepen. Zij hebben niet geparticipeerd in de gegeven beoordelingen en waren toevoerder.