

Van: Stuurgroep Zichtbare Zorg GGZ

Datum: 17 december 2010

Betreft: Aanbiedingsbrief bij het evaluatierapport Basisset prestatie-indicatoren GGZ en VZ verslagjaar 2009 (versie 11 november 2010)

De stuurgroep Zichtbare Zorg GGZ biedt u het evaluatierapport Basisset prestatie-indicatoren Geestelijke Gezondheidszorg en Verslavingszorg over verslagjaar 2009 aan. De stuurgroep Zichtbare Zorg GGZ bestaat uit de volgende branche-/koepelorganisaties: ZN, LPGGz, IGZ, VWS, NVvP, Platform MEER GGZ, NIP, NVP, GGZ Nederland en V&VN. Deze stuurgroep bepaalt het beleid omtrent de prestatie-indicatoren binnen de GGZ. Jaarlijks worden de aangeleverde data op prestatie-indicatoren geëvalueerd.

Voor u ligt het evaluatierapport over verslagjaar 2009, de vierde landelijke uitvraag. In bijlage J staan per indicator de scores van de verschillende uitgevoerde analyses samengevat. Uit deze resultaten blijkt dat het nog steeds moeilijk is om valide, betrouwbaar en vergelijkbaar te meten dan wel gewenste gegevens aan te leveren. Redenen zijn als volgt:

- Onvolledigheid van aanlevering. De mate waarin de prestatie-indicatoren worden aangeleverd, kan voor nagenoeg alle indicatoren nog sterk worden verbeterd; in veel gevallen wordt als reden voor het niet aanleveren van indicatoren aangegeven dat gegevens niet geregistreerd zijn of dat de informatie lastig uit de registratie te halen is;
- Gebruik meerdere meetinstrumenten. Voor de indicatoren zijn nog verschillende meetinstrumenten inclusief eigen meetinstrumenten gebruikt en niet alle CQ-resultaten zijn gecorrigeerd voor verschillen in casemix. Dit gaat ten koste van de vergelijkbaarheid van de data;
- Kleine aantallen. Veel van de GGZ-instellingen die gegevens aangeleverd hebben zijn beperkt in omvang. Dit betekent dat er voor een beperkt aantal cliënten waarnemingen zijn. Dit gaat ten koste van de statistische betrouwbaarheid; hoe lager het aantal waarnemingen des te meer berust de gemeten indicatorwaarde op toeval.

Bovenstaande maakt dat er nog niet voor alle indicatoren betrouwbare vergelijkingen tussen instellingen kunnen worden gemaakt. Het is wel mogelijk dat instellingen zichzelf vergelijken over de jaren heen waardoor de gegevens handvatten opleveren om de kwaliteit van de zorg intern te verbeteren.

Om de validiteit, betrouwbaarheid en vergelijkbaarheid te verhogen en de registratielast te verlagen zijn er voor de prestatie-indicatoren voor verslagjaar 2010 en 2011 een aantal verbeterlagen doorgevoerd. Zo is onder andere de basisset gereduceerd tot een kernset, wordt de CQ-index Kortdurend ambulante GGZ en VZ nu verplicht, zijn er nieuwe CQ-indexen voor langdurige en klinische GGZ en wordt er een start gemaakt om de effectiviteitindicatoren via ROM-gegevens in te vullen, zodat op termijn een benchmark mogelijk wordt.

De stuurgroep Zichtbare Zorg GGZ doet de volgende aanbevelingen aan instellingen die de metingen moeten uitvoeren en aan externe partijen die deze gegevens interpreteren:

- Maak een analyse per instelling van de gegevens van afgelopen vier jaar en ga hierover het gesprek aan binnen de zorginstelling en met externe partijen. De instelling stelt vervolgens verbeterplannen op en neemt dit op in het jaarplan;

- Vermijd indien mogelijk het gebruik van eigen instrumenten, aangezien de indicatorwaarden dan niet vergeleken kunnen worden met die van andere instellingen;
- Verbeter het registratieproces (o.a. ROM, CQi) binnen de instelling om als eerste die indicatoren die wettelijk verplicht zijn volledig aan te leveren.

Meer informatie over het project Zichtbare Zorg GGZ kunt u vinden op de website www.zichtbarezorg.nl.



Evaluatierapport Basisset prestatie-indicatoren Geestelijke Gezondheidszorg

Data aanlevering verslagjaar 2009

Versie Stuurgroep 11 november 2010
Data-analyse door advies- en onderzoeksbureau Significant
Uitgave van Bureau Zichtbare Zorg Geestelijke Gezondheidszorg

Inhoudsopgave

1 Inleiding	6
<i>Achtergrond</i>	6
<i>Doelgroep</i>	6
<i>Proces</i>	6
<i>Opzet rapport sluit aan op aanpak Zichtbare Zorg-breed</i>	7
<i>Leeswijzer</i>	7
<i>Aanvullende informatie</i>	7
2 Resultaten en conclusies	9
2.1 Onderzoeksstappen.....	9
2.1.1 <i>Datakwaliteit</i>	9
2.1.2 <i>Uitgevoerde analyses</i>	9
<i>Volledigheid</i>	10
<i>Betrouwbaarheid van het registratieproces</i>	10
<i>Statistisch betrouwbare vergelijkbaarheid</i>	10
<i>Onderscheidend vermogen</i>	10
<i>Correlaties tussen de meetjaren</i>	11
<i>Validiteit</i>	11
<i>Populatievergelijkbaarheid</i>	11
2.1.3 <i>Beperkingen van het databestand</i>	11
<i>Resultaten per instellingstype zijn niet te onderscheiden</i>	11
<i>Resultaten van een meting met een eigen meetinstrument zijn onderling niet vergelijkbaar</i>	12
<i>CQ-index resultaten zijn onderverdeeld naar 'zelf aangeleverd' en 'via Trusted Third Party' aangeleverd</i>	12
<i>De schoningsvragen die Zichtbare Zorg aan de GGZ-instellingen heeft gesteld zijn 'attenderingsvragen'</i>	12
2.2 Volledigheid	13
<i>Aanpak</i>	13
<i>Resultaten</i>	13
<i>Conclusies</i>	17
2.4 Betrouwbaarheid	19
2.4.1 <i>Statistisch betrouwbare vergelijkbaarheid</i>	19
<i>Aanpak</i>	19
<i>Resultaten</i>	21
<i>Conclusies</i>	21
2.4.2 <i>Onderscheidend vermogen</i>	22
<i>Aanpak</i>	22
<i>Resultaten</i>	23
<i>Conclusies</i>	24
2.4.3 <i>Correlaties tussen de meetjaren</i>	25
<i>Aanpak</i>	25
<i>Resultaten</i>	26
<i>Conclusies</i>	29
2.4 Validiteit en populatievergelijkbaarheid	31
<i>Aanpak</i>	31
<i>Resultaten</i>	32
<i>Conclusies</i>	35

2.5 Conclusie.....	36
3 Aanbevelingen.....	37
3.1 Procesmatige verbeterpunten.....	37
3.2 Inhoudelijke verbeterpunten	37
3.3 Verbeterpunten ten aanzien van de portal en dataset.....	38
Bijlage A	Kenmerken instellingen
Bijlage B	Landelijke spiegelinformatie
Bijlage C	Samenvattende indicatorentabel
Bijlage D	Toelichting bij de figuren en tabellen
Bijlage E	Aanvullende resultaten volledigheid
Bijlage F	Overzichtstabellen resultaten
Bijlage G	Signaalvlaggen voor statistische betrouwbaarheid
Bijlage H	Typering van de instellingen
Bijlage I	Formulering van een vraag in het eigen meetinstrument
Bijlage J	Samenvattende tabellen van de uitgevoerde analyses
Bijlage K	Validiteit

1 Inleiding

Achtergrond

Kwalitatief goede zorg is zorg die patiëntgericht, effectief, veilig en op tijd is en die is toegesneden op de behoeften van de individuele patiënt en cliënt. Het zichtbaar maken van de kwaliteit van de zorg en van verschillen in kwaliteit is essentieel voor de werking van het nieuwe zorgstelsel. De kern van het nieuwe zorgstelsel is dat de patiënt en consument in staat worden gesteld om geïnformeerd te kiezen; kiezen op het niveau van de instelling en aandoening. Daarvoor is vergelijkbare en betrouwbare informatie nodig over de prestaties van zorgaanbieders.

Het programma Zichtbare Zorg ondersteunt zorgaanbieders van Geestelijke Gezondheidszorg bij het transparant maken van de kwaliteit van de geleverde zorg. Aan het programma nemen zowel zorgaanbieders als zogenaamde 'vragende partijen' deel, zoals zorgverzekeraars, consumenten en patiënten, en ook de Inspectie voor de Gezondheidszorg.

Het voorliggende evaluatierapport draagt bij aan het transparant maken van de zorg. De rapportage bevat een analyse van de gegevens die Geestelijke Gezondheidszorg (GGZ)-instellingen dit jaar bij Zichtbare Zorg hebben aangeleverd over de kwaliteit van zorg. Het rapport heeft betrekking op zorg die is geleverd in 2009.

Het rapport beschrijft de volledigheid, betrouwbaarheid, validiteit en populatievergelijkbaarheid van de aangeleverde data. Ook biedt het rapport een overzicht van de resultaten: een grafiek of tabel met de indicatorwaarden.

Doelgroep

Dit evaluatierapport is bedoeld voor de volgende gebruikers:

1. Gebruikers die van de data informatie maken voor hun specifieke doel en/of doelgroep, zoals beleidsmedewerkers/projectleiders bij patiëntenorganisaties, adviseurs zorginkoop bij zorgverzekeraars en kwaliteitsmanagers in de instellingen.
2. De verschillende partijen betrokken bij de organisatie van Zichtbare Zorg GGZ: de stuurgroep, projectgroep en Bureau Zichtbare Zorg. Zij krijgen met deze rapportage een totaalbeeld van de indicatoren. Het plaatst keuzes voor doorontwikkeling in perspectief.
3. Instellingen kunnen ook gebruikmaken van deze rapportage. Zij ontvangen echter ook een spiegelrapportage waarin hun eigen scores ten opzichte van het landelijk beeld zijn weergegeven. Op deze manier wil Zichtbare Zorg instellingen meer inzicht bieden in hun eigen organisatie, input leveren ten behoeve van de zorginkoop en bijdragen aan een sterkere positionering in de markt.
4. De individuele burger of kiezende consument is nadrukkelijk niet de doelgroep van dit evaluatierapport.

Proces

Alle Jaarverslag Maatschappelijke Verantwoording (JMV)-plichtige GGZ-instellingen hebben dit jaar voor de vierde keer op rij gegevens aangeleverd over de kwaliteit van de geleverde zorg. De gegevens zijn van 15 februari tot 1 mei 2010 door 178 GGZ-instellingen aangeleverd en hebben betrekking op het verslagjaar 2009.

Vervolgens werden de gegevens geschoond: bij onwaarschijnlijke waarden vroeg Zichtbare Zorg de GGZ-instellingen de aangeleverde gegevens te checken en waar nodig te wijzigen. De definitieve gegevens zijn geaccordeerd door de instellingen.

In september en oktober zijn de gegevens geanalyseerd door Zichtbare Zorg. Het voor u liggende evaluatierapport is opgesteld, en de instellingen ontvingen online een spiegelrapportage. Vanaf 1 juli 2010 zijn de gegevens van de GGZ-instellingen over het verslagjaar 2009 opvraagbaar bij het JMV. Hierbij dient opgemerkt te worden dat de gegevens van drie instellingen niet meegenomen konden worden in dit rapport, omdat deze gegevens in een later stadium beschikbaar kwamen. De gegevens van die drie instellingen zijn medio oktober 2010 aan het JMV toegevoegd.

Het voor u liggende evaluatierapport is vanaf november 2010 openbaar beschikbaar op de website van Zichtbare Zorg.

Opzet rapport sluit aan op aanpak Zichtbare Zorg-breed

In dit rapport wordt een eerste aanzet gegeven om de kwaliteit van de, door de GGZ-instellingen aangeleverde, data vast te stellen op een Zichtbare Zorg-brede wijze. Deze wijze heeft Zichtbare Zorg de afgelopen jaren ontwikkeld en is nog in ontwikkeling. De kwaliteit van de data wordt middels 'signaalvlaggen' voor de volgende elementen bepaald: validiteit, betrouwbaarheid van het registratieproces, statistisch betrouwbare vergelijkbaarheid en populatievergelijkbaarheid. Door het toekennen van signaalvlaggen, in rood, oranje of groen, wordt in een oogopslag duidelijk wat de kwaliteit is van de aangeleverde data. De vlaggen zeggen niets over de kwaliteit van de geleverde zorg. In dit rapport zijn de signaalvlaggen berekend voor de statistisch betrouwbare vergelijkbaarheid. De vlaggen geven per indicator aan in hoeverre het mogelijk is om statistisch betrouwbare uitspraken te doen over verschillen in resultaten tussen verschillende instellingen.

Aan de andere kwaliteitsparameters die in dit rapport zijn berekend zijn geen vlaggen toegekend, omdat dat op basis van de beschikbare informatie en gekozen aanpak niet mogelijk is.

Leeswijzer

De rapportage bestaat uit twee hoofdstukken:

- Hoofdstuk 2 is een uitgebreide beschrijving en bevat conclusies van de resultaten van de analyses van de volledigheid, betrouwbaarheid, validiteit en populatievergelijkbaarheid;
- In hoofdstuk 3 worden aanbevelingen gedaan.

In de bijlagen staat de volgende achtergrondinformatie:

- Bijlage A: Kenmerken instellingen
- Bijlage B: Landelijke spiegelinformatie
- Bijlage C: Samenvattende indicatorentabel
- Bijlage D Toelichting bij de figuren en tabellen
- Bijlage E: Aanvullende resultaten volledigheid
- Bijlage F: Overzichtstabellen resultaten
- Bijlage G: Signaalvlaggen voor statistische betrouwbaarheid
- Bijlage H: Typering van de instellingen
- Bijlage I: Formulering van een vraag in het eigen meetinstrument
- Bijlage J: Samenvattende tabellen van de uitgevoerde analyses

Aanvullende informatie

De rapportage is zo beknopt als mogelijk gehouden en focust op de evaluatie van de aangeleverde gegevens over het verslagjaar 2009 over de Basisset Prestatie Indicatoren

GGZ. Instellingen konden tegelijk met de Basisset GGZ, nog een andere indicatorset aanleveren: die voor Forensische Psychiatrie (FP). Voor FP was het mogelijk om de indicatorset over 2007/2008 (acht instellingen) of die van 2009/2010 (dertien instellingen) aan te leveren.

De landelijke spiegelinformatie van de FP-set wordt apart gerapporteerd. De set is niet geëvalueerd op volledigheid, betrouwbaarheid en validiteit, omdat het aantal aanleverende instellingen van deze sets voor veel uit te voeren analyses te laag is. Meer informatie staat op de website van Zichtbare Zorg (www.zichtbarezorg.nl).

2 Resultaten en conclusies

In dit hoofdstuk staan de onderzoeksstappen kort beschreven. Vervolgens worden de aanpak, resultaten en conclusies van de verschillende analyses behandeld.

2.1 Onderzoeksstappen

De gegevens die de instellingen via de webportal van Zichtbare Zorg hebben aangeleverd, zijn in één databestand opgenomen. Dit databestand is gecontroleerd en verrijkt met gegevens uit het voorgaande verslagjaar. In deze paragraaf zijn de uitgevoerde controles en analyses beschreven. Tevens bespreken we een aantal beperkingen van het databestand.

2.1.1 Datakwaliteit

Voordat de analyses, beschreven in de volgende paragraaf, zijn uitgevoerd, zijn de volgende controles op de datasets uitgevoerd:

- Bij de koppeling van de data over het verslagjaar 2008 aan het verslagjaar 2009, is op aanwijzen van Zichtbare Zorg GGZ vastgesteld welke indicatoren op inhoudelijke grond met elkaar konden worden vergeleken.
- Bij de koppeling van de data over het verslagjaar 2008 aan het verslagjaar 2009, zijn de instellingen op KvK-nummer aan elkaar gekoppeld. In een aantal gevallen waar dit niet mogelijk was, is op aanwijzen van Zichtbare Zorg alsnog een koppeling gemaakt.
- De analyses zijn uitgevoerd en ter verificatie steekproefsgewijs gecontroleerd.

Naar aanleiding van verzoeken uit het veld is bekeken of het mogelijk was om in de analyses onderscheid te maken naar type instelling. Dit is op methodologische gronden niet mogelijk, zie ook de paragraaf over de beperkingen van de dataset. In bijlage H zijn de aanleverende instellingen opgenomen, inclusief de type instelling waartoe zij zelf hebben aangevoerd toe te behoren.

2.1.2 Uitgevoerde analyses

Om de kwaliteit van de aangeleverde data vast te stellen zijn er zes analyses uitgevoerd. Onderstaande tabel geeft aan welke analyses in welke paragraaf staan, en of de analyse onderdeel uitmaakt van de signaalvlaggenmethodiek van Zichtbare Zorg, zoals deze op dit moment is. De signaalvlaggenmethodiek is met een expertgroep ontwikkeld in het kader van Zichtbare Zorg Ziekenhuizen, en is nog in ontwikkeling. In dit rapport is een eerste aanzet gegeven van toepassing van de methodiek op de GGZ-dataset. Onder de tabel is een korte toelichting gegeven op iedere analyse. In de paragrafen waar de analyses zijn uitgevoerd, is de aanpak verder toegelicht.

Analyse	Paragraaf	Onderdeel signaalvlaggen methodiek
Volledigheid	2.2	nee
Betrouwbaarheid registratieproces	-	ja
Statistisch betrouwbare vergelijkbaarheid	2.3.1	ja
Onderscheidend vermogen	2.3.2	nee
Correlaties tussen de meetjaren	2.3.3	nee
Validiteit	2.4.1	ja
Populatievergelijkbaarheid	2.4.2	ja

Volledigheid

De mate waarin GGZ-instellingen de basisset prestatie-indicatoren volledig hebben aangeleverd is relevant, omdat instellingen vaak niet volledig alle indicatoren aanleveren. Het niet volledig aanleveren van gegevens is voor een deel onderdeel van de indicatorset, omdat sommige indicatoren verplicht zijn, terwijl andere facultatief zijn om aan te leveren.

Betrouwbaarheid van het registratieproces

De betrouwbaarheid van het registratieproces, onderdeel van de signaalvlaggenmethodiek, kan op basis van het databestand niet worden bepaald. Hieronder wordt toegelicht welke gegevens er nodig zijn om de betrouwbaarheid van het registratieproces vast te stellen, en welke stappen er zijn ondernomen om volgend jaar de signaalvlag wel te berekenen.

Om een goed beeld te krijgen van de betrouwbaarheid van het registratieproces, moet er antwoord worden gegeven op de volgende vragen:

- Is het gevraagde verslagjaar toegepast?
- Is de populatie vastgesteld op de voorgeschreven manier?
- Is het mogelijk geweest om de gevraagde in- en exclusiecriteria te hanteren?
- Is er sprake van schriftelijke goedkeuring voor het aanleveren van de gegevens?
- Worden de gegevens intern gebruikt als informatie voor het kwaliteitsbeleid?

Om betrouwbare antwoorden te krijgen op deze vragen, is het noodzakelijk deze vragen direct te stellen op het moment dat de gegevens door de instellingen worden aangeleverd. Op die manier krijg je een nauwkeurig en voor alle instellingen representatief beeld van de betrouwbaarheid van het registratieproces. Bovenstaande vragen zijn echter niet direct tijdens het aanleveren van de gegevens aan de instellingen gevraagd. Uit de open opmerkingen die de instellingen per indicator konden maken, staan diverse opmerkingen die mogelijk over de betrouwbaarheid van het registratieproces gaan. Op basis van de open opmerkingen is het niet mogelijk om per indicator algemeen geldende conclusies te trekken over de betrouwbaarheid van het registratieproces.

Daarom is dit jaar een eerste aanzet gegeven om bovenstaande vragen te stellen die nodig zijn om de betrouwbaarheid van het registratieproces te bepalen. De vragen (ook wel 'verificatievragen') zullen bij de eerst volgende uitvraag worden toegevoegd aan de uitvraag van de indicatoren.

Statistisch betrouwbare vergelijkbaarheid

Dit jaar is voor de basisset prestatie-indicatoren GGZ voor het eerst de statistisch betrouwbare vergelijkbaarheid bepaald aan de hand van signaalvlaggen. De signaalvlag zegt iets over de nauwkeurigheid van de aangeleverde waarde op zichzelf en in vergelijking met de gemiddelde waarde over alle instellingen (het landelijke gemiddelde). Dit is in principe niet door een instelling te beïnvloeden. Daarom rapporteren we deze vlaggen, analoog aan andere Zichtbare Zorg-sectoren, alleen op landelijk niveau, en niet op instellingsniveau.

Onderscheidend vermogen

Indicatoren worden voor gebruikers interessanter als de resultaten significante verschillen tussen de zorgaanbieders laten zien. De analyse naar het onderscheid vermogen stelt vast in welke mate er significante verschillen tussen GGZ-instellingen bestaan. Alleen voor die indicatoren waarbij het statistisch gezien verantwoord is, is deze analyse uitgevoerd.

Correlaties tussen de meetjaren

De betrouwbaarheid van de indicatoren is bekeken om de samenhang van de resultaten tussen beide meetjaren vast te stellen. Een van de verklaringen voor een hoge samenhang is dat de indicator betrouwbaar is.

Validiteit

De validiteit van de indicator geeft aan in hoeverre de indicator meet wat hij hoort te meten. Dit element wordt op het niveau van de indicator vastgesteld, en niet bijvoorbeeld per instelling. In evaluatiebijeenkomsten die met GGZ-instellingen zijn gehouden, is dit vastgesteld.

Populatievergelijkbaarheid

De populatievergelijkbaarheid zegt iets over de mate waarin indicatorwaarden tussen verschillende instellingen goed met elkaar te vergelijken zijn. Voor sommige indicatoren geldt dat kenmerken van de instelling (denk aan patiëntkenmerken) sterke invloed hebben op de indicatorwaarde. In de evaluatiebijeenkomsten hebben de GGZ-instellingen aangegeven voor welke indicatoren dit met name geldt.

2.1.3 Beperkingen van het databestand

Resultaten per instellingstype zijn niet te onderscheiden

In de gehouden evaluatiebijeenkomsten is door een aantal GGZ-instellingen benadrukt dat voorzichtigheid is geboden bij het maken van een vergelijking tussen verschillende instellingen. Tevens bestaat de wens om in het rapporteren van de resultaten juist onderscheid te maken naar type instelling. Het maken van een onderscheid in resultaten tussen de verschillende typen instellingen is, op basis van de gegevens over het verslagjaar 2009, niet mogelijk. De reden hiervan wordt hieronder toegelicht.

Over het verslagjaar 2009 hebben instellingen zelf aangegeven bij welk type instelling ze zichzelf rekenen. Instellingen konden daarbij meerdere typen aangeven. Vervolgens hebben de instellingen alle indicatoren die voor de instelling van toepassing waren, kunnen aanleveren. Of een indicator van toepassing was voor een instelling, is niet bepaald aan de hand van de keuze voor type instelling, maar aan de hand van andere kenmerken van een instelling. Bijvoorbeeld of een instelling langdurige zorg levert of niet. De indicatorwaarden die de instelling leverde, zijn bovendien niet uitgesplitst naar type instelling uitgevraagd. Hierdoor is het niet mogelijk om de waarden van de indicatoren per type instelling te bepalen. In bijlage H is een overzicht opgenomen van alle instellingen en de typen die zij zelf hebben aangegeven te zijn.

Stel een GGZ-instelling heeft bijvoorbeeld aangegeven tot de volgende typen instellingen te horen: Geïntegreerde GGZ-instelling, Verslavingszorg, Ambulante GGZ, RIBW, Forensische psychiatrie en SGLVG. De waarde die die instelling heeft aangeleverd voor bijvoorbeeld indicator '1.3 Percentage cliënten met de hoofddiagnose schizofrenie en andere psychotisch stoornis of bipolaire stoornis waarbij dagelijks functioneren bij voor- en nameting is verbeterd' kan niet per type instelling worden uitgesplitst, omdat de waarde niet per type is aangeleverd. Ook is het niet correct om de waarde die de instelling heeft aangegeven bij slechts een van de aangegeven typen toe te kennen, omdat het niet bekend is op basis van welke cliëntenmix de instelling tot die waarde is gekomen.

Resultaten van een meting met een eigen meetinstrument zijn onderling niet vergelijkbaar

De resultaten van de metingen die instellingen hebben uitgevoerd met een eigen meetinstrument kunnen niet in één figuur staan. De resultaten verkregen met een eigen meetinstrument hebben het kenmerk dat de vraag die wordt gesteld per instelling verschilt. Dit is ook zichtbaar in de dataset, want instellingen hebben aangegeven op welke manier de vraag is geformuleerd. Ter illustratie zijn in Bijlage I de formuleringen van een vraag opgenomen. Naast de vragen die bij metingen met het eigen meetinstrument pluriform zijn geformuleerd, kunnen de antwoordschalen ook divers zijn opgesteld. Zo kan de tevredenheid bijvoorbeeld op een ja/nee-schaal worden uitgevraagd, of op een vijf-puntsschaal, of als een rapportcijfer. Vergelijkbaarheid van resultaten verkregen met een eigen meetinstrument is daarom niet zinvol.

CQ-index resultaten zijn onderverdeeld naar 'zelf aangeleverd' en 'via Trusted Third Party' aangeleverd

De resultaten die door de instelling zelf zijn aangeleverd, zijn niet gecorrigeerd voor verschillen in casemix. De resultaten die zijn aangeleverd via de 'Trusted Third Party' zijn wel gecorrigeerd voor verschillen in casemix. Over het verslagjaar 2009 zijn er zeventien instellingen die dat hebben gedaan.

De schoningsvragen die Zichtbare Zorg aan de GGZ-instellingen heeft gesteld zijn 'attenderingsvragen'

De GGZ-instellingen hebben van 15 februari tot 1 mei 2010 gegevens aangeleverd bij Zichtbare Zorg. Vervolgens heeft Zichtbare Zorg 'schooningsvragen' gesteld aan de indicatoren. Schooningsvragen zijn gesteld om instellingen te attenderen op opmerkelijke waarden, om de instellingen de mogelijkheid te geven die waarden te controleren, en indien nodig aan te passen. Het wel of niet aanpassen van de aangeleverde waarde door instellingen heeft geen verband met de betrouwbaarheid van de waarde.

2.2 Volledigheid

Aanpak

De volledigheid van de aangeleverde data is vastgesteld per (sub)indicator en per instelling. Per indicator is bepaald:

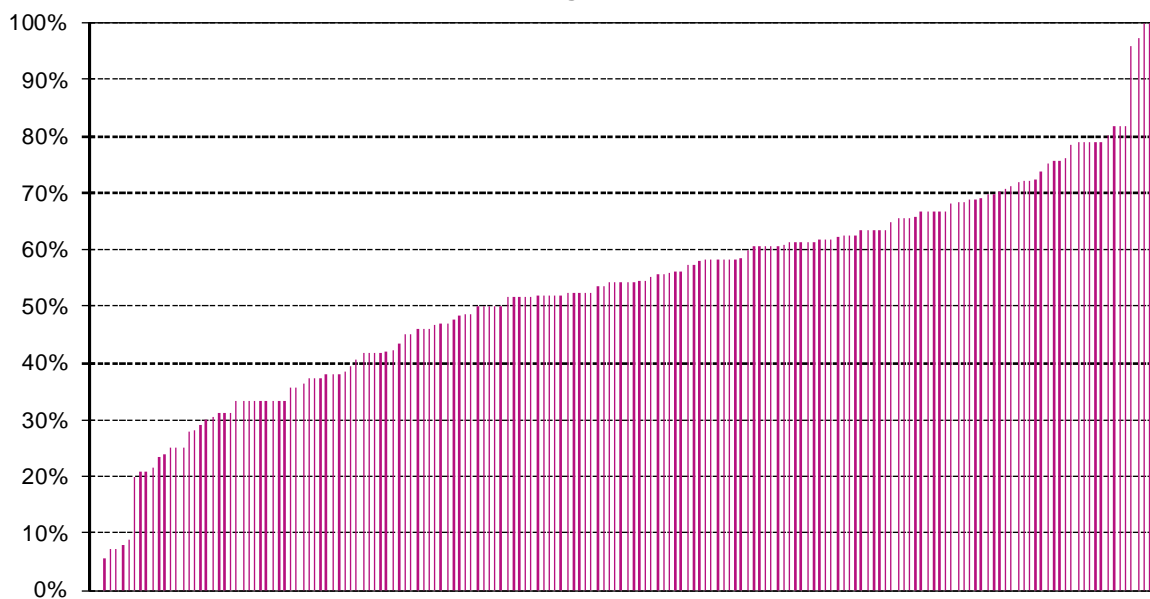
1. Welke indicatoren voor instellingen van toepassing zijn;
2. Welke instellingen zelf aangeven geen gegevens voor de indicator te kunnen aanleveren;
3. Welke instellingen daadwerkelijk een waarde hebben aangeleverd;
4. Per indicator is het percentage berekend van het aantal daadwerkelijk aanleverende instellingen, ten opzichte van het totaal aantal instellingen waarop de indicator van toepassing is.

Resultaten

De Basis Prestatie Indicatorenset GGZ is over het verslagjaar 2009 aangeleverd door 178 instellingen. Afhankelijk van de kenmerken van de instellingen, die de instellingen zelf hebben aangegeven, varieert het aantal prestatie-indicatoren dat van toepassing is op de instellingen tussen 9 en 38. Er zijn enkele instellingen die alle van toepassing zijnde indicatoren hebben aangeleverd. Uit de onderstaande figuur blijkt dat er een grote variabiliteit is tussen instellingen in de mate van volledigheid van aanlevering. Er zijn twee instellingen die aangegeven hebben de indicatorset aan te leveren, maar uiteindelijk bij geen enkele indicator waarden hebben ingevuld. 9 instellingen hebben 80% of meer van de op de instelling van toepassing zijnde indicatoren aangeleverd.

Overigens is het niet het geval dat instellingen waarop slechts een klein aantal indicatoren van toepassing is, in verhouding meer volledig aanleveren dan instellingen die (bijna) de gehele set dienen aan te leveren. Tevens is er geen relatie te zien tussen de omvang van de instelling (in aantal cliënten dat is behandeld) en de mate van volledigheid waarin is aangeleverd. Figuren die dit illustreren staan in bijlage E met Aanvullende resultaten volledigheid.

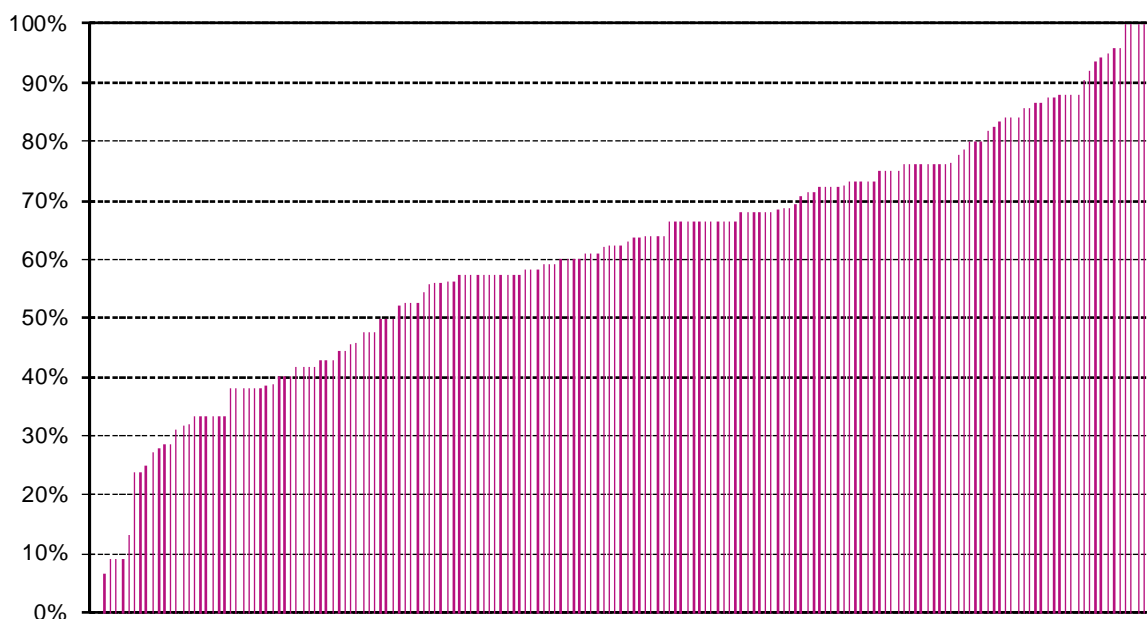
Per instelling het percentage van de van toepassing zijnde indicatoren dat is aangeleverd



Instellingen gerangschikt naar percentage aangeleverd
(percentages: gemiddelde = 53%, mediaan = 54%, min = 0%, max = 100%)
(aantal indicatoren van toepassing: gemiddelde = 25, mediaan = 24, min = 9, max = 38)

Over het verslagjaar 2009 was het niet voor alle prestatie-indicatoren verplicht om gegevens aan te leveren. Voor 13 van de 38 indicatoren was aanlevering facultatief. Hoewel verplichte indicatoren volledig ingevuld zouden moeten zijn, laat de onderstaande figuur zien dat gemiddeld 61% van de verplichte prestatie-indicatoren aangeleverd is. Er zijn 48 instellingen die voor minder dan de helft van de verplichte indicatoren gegevens hebben aangeleverd en 32 instellingen die meer dan 80% van de verplichte indicatoren hebben aangeleverd.

Per instelling het percentage van de verplichte indicatoren dat is aangeleverd



Instellingen gerangschikt naar percentage aangeleverd
(percentages: gemiddelde = 61%, mediaan = 63%, min = 0%, max = 100%)
(aantal verplichte indicatoren: gemiddelde = 19, mediaan = 21, min = 7, max = 25)

In de tabel op de volgende pagina is per indicator weergegeven op hoeveel instellingen de indicator van toepassing is en hoeveel instellingen daadwerkelijk gegevens hebben aangeleverd.

Een aantal indicatoren is opvallend volledig aangeleverd. Dit geldt met name voor de prestatie-indicatoren uit het onderdeel 'cliëntgerichtheid', waarbinnen acht van de dertien (sub)indicatoren is aangeleverd. Dit deed meer dan 80% van de instellingen waarop de indicatoren van toepassing is. Echter, voor een groter deel van de indicatoren is door minder dan de helft van de instellingen gegevens aangeleverd. Dit is voor de prestatie-indicatoren uit het onderdeel 'veiligheid' niet verwonderlijk, aangezien aanlevering van deze indicatoren facultatief was. Echter, ook een groot deel van de verplichte prestatie-indicatoren uit het onderdeel 'effectiviteit' is minder volledig aangeleverd, met uitzondering van de indicatoren 1.2a 'Oordeel van de cliënt over de verandering ernst problematiek' en 1.11 'Tijdig contact na ontslag uit kliniek'. Met betrekking tot de (sub)prestatie-indicatoren 1.4 (over de kwaliteit van leven) en 1.5 (over rehabilitatie) geeft 65% tot 80% van de instellingen aan dat er binnen de instelling geen gegevens beschikbaar zijn over het betreffende onderwerp.

Als een indicator is uitgevraagd onder verschillende diagnosegroepen, zijn er slechts geringe verschillen in de volledigheid van de aanlevering tussen de verschillende diagnosegroepen. Deze staan in bijlage E met Aanvullende resultaten volledigheid.

Er zijn zeven prestatie-indicatoren die betrekking hebben op cliëntervaringen waarbij de instellingen de keuze hadden tussen de volgende drie meetinstrumenten voor het aanleveren van gegevens: de Cliëntthermometer, een eigen meetinstrument of de CO-index. Vooraf hebben instellingen aangegeven met welk(e) meetinstrument(en) zij de cliëntervaringen gemeten hadden. Het grootste deel van de instellingen heeft de Cliëntthermometer gebruikt voor het meten van cliëntervaring. De gegevens voor de

betreffende indicatoren dan ook het vaakst zijn aangeleverd op basis van de cliëntthermometer. Maar ook relatief gezien blijkt dat als gemeten is met de Cliëntthermometer, de indicatoren vaker zijn ingevuld vergeleken met een eigen meetinstrument of de CQ-index. Zie ook bijlage E met Aanvullende resultaten volledigheid.

Volledigheid van de aanlevering per indicator			
Indicator	Aantal instellingen waarop indicator van toepassing is	Aantal instellingen dat heeft aangeleverd	Percentage van de instellingen waarop de indicator van toepassing is dat heeft aangeleverd
Deel 1 Effectiviteit			
1.1 Bereik preventieactiviteiten	178	60	34%
1.2a Oordeel cliënt verandering ernst problematiek	149	134	90%
1.2b Systematisch meten verandering ernst problematiek	156	61	39%
1.2c Gebruikte instrumenten	156	74	47%
1.2d Verandering ernst problematiek ¹	156	59	38%
1.3 Verandering in het dagelijks functioneren van de cliënt ¹	123	39	32%
1.4a Systematisch meten van verandering in ervaren kwaliteit	138	21	15%
1.4b Gebruikte instrumenten	138	48	35%
1.4c Verandering in de ervaren kwaliteit van leven	138	18	13%
1.5a Percentage metingen naar onvervulde zorgbehoefte	111	16	14%
1.5b Gebruikte instrumenten	111	36	32%
1.7 Drop-out	113	72	64%
1.8a Standaard procedure voor somatische screening	101	56	55%
1.8b Mate waarin screening op somatische aandoeningen plaatsvindt	135	77	57%
1.11 Tijdig contact na ontslag uit kliniek	65	54	83%
Deel 2 Veiligheid			
2.1a Beschikbaarheid medicatie-overzicht ^F	156	32	21%
2.1b Combinatie SSRI NSAID met PPI ^F	156	13	8%
2.1c Rationeel voorschrijven benzodiazepines ^F	156	15	10%
2.2 Informatie over bijwerkingen medicijnen (CQ-index) ^F	30	23	77%
2.3 Onveiligheid door ontbreken van informatie ^F	54	22	41%
2.4a Aantal insluitingen / toepassingen van dwangmedicatie ^F	66	17	26%
2.4b Gemiddelde duur insluiting / dwangmedicatie ^F	66	15	23%
2.4c Evaluatie met client ^F	66	7	11%
2.4d Aanwezigheid van preventieprogramma's ^F	66	30	45%
2.6a Ervaren veiligheid cliënten - CQ-index ^F	15	4	27%
Deel 3 Cliëntgerichtheid			
3.1a Gemiddelde aanmeldingswachttijd	178	155	87%
3.1b Gemiddelde wachttijd tot behandeling	178	145	81%
3.1c Percentage wachttijd korter dan x dagen	178	149	84%
3.3 Informed consent	166	132	80%
3.4 Keuzevrijheid	171	149	87%
3.5 Vervulling zorgwensen	171	151	88%
3.6a Cliëntoordeel evaluatie of bijstelling behandel- of begeleidingsplan ^F	27	11	41%
3.6b Percentage geëvalueerde of bijgestelde behandel- of begeleidingsplannen	142	99	70%
3.7 Goede samenwerking en afstemming met ketenpartners ^F	33	26	79%
3.8a Cliëntervaring zinvolle dagbesteding ^F	27	5	19%
3.9 Cliëntoordeel woon- en leefomstandigheden	171	50	29%
3.10 Adequate bejegening door hulpverleners	171	154	90%
3.11 Adequate informatieverstrekking door hulpverlener	171	150	88%

¹Bij de indicatoren 1.2d en 1.3 wordt gevraagd naar zowel verbetering als stabilisatie. Beide subonderdelen worden door dezelfde instellingen ingevuld, daarom zijn ze bij de bepaling van de volledigheid samengenomen

^F Voor deze indicatoren is de aanlevering over het verslagjaar 2009 facultatief

Er zijn verschillende redenen voor het niet-aanleveren van indicatoren:

- *Geen gegevens beschikbaar of niet geregistreerd.* Een belangrijke oorzaak voor het niet-aanleveren van gegevens is dat er binnen de instelling geen gegevens beschikbaar zijn over de betreffende prestatie-indicator. Instellingen vinken gemiddeld bij 17 indicatoren het vakje "geen gegevens beschikbaar" aan. Dit vakje was niet op het niveau van de subindicatoren aan te vinken, wat logischerwijs wel wenselijk was. Dit heeft ertoe geleid dat instellingen soms waarden moesten aanleveren, terwijl zij dat niet konden. Hierdoor zijn mogelijk nullen ingevoerd. Het is niet vast te stellen om hoeveel gevallen het gaat. In de open opmerkingsvelden benoemen instellingen regelmatig dat de gevraagde gegevens niet geregistreerd zijn, of niet op systematische wijze gemeten of geregistreerd worden, waardoor het lastig is de gegevens te achterhalen. Dit is waarschijnlijk ook de reden dat bij de indicatoren 1.4 en 1.5 met betrekking tot de effectiviteit van de behandeling slechts een zeer klein aantal instellingen heeft aangeleverd. Bij vijf van de acht indicatoren uit het onderdeel 'effectiviteit' wordt beperkte registratie regelmatig genoemd in de opmerkingsvelden. Instellingen konden aangeven dat zij geen gegevens beschikbaar hadden om indicatoren aan te leveren. Daarnaast wordt er bij enkele instellingen niet op het niveau van diagnosegroepen geregistreerd, waardoor aanlevering van de indicatoren die opgesplitst zijn in diagnosegroepen niet mogelijk is.
- *Indicator is niet van toepassing op de instelling.* Ondanks dat er een routing in de webportal zit, komt het nog regelmatig voor dat een indicator niet van toepassing is op een instelling, omdat de instelling niet met de betreffende zorg of situatie te maken heeft, zoals dwangverpleging, drop-out, benzodiapines en wachtlijsten. Ook zijn er instellingen die alleen begeleiding bieden en geen behandeling, waardoor zij indicatoren over de effectiviteit van de behandeling niet kunnen aanleveren.
- *Facultatieve aanlevering.* De indicatoren waarbij aanlevering facultatief is, worden door zeer weinig instellingen aangeleverd. In de open opmerkingsvelden bij de facultatieve prestatie-indicatoren wordt regelmatig benoemd dat de indicator niet wordt aangeleverd juist omdat deze facultatief is. Dit duidt erop dat instellingen nauwelijks een prikkel hebben om gegevens aan te leveren zolang aanlevering niet verplicht gesteld is.
- *Overige redenen.* Enkele instellingen zijn pas net van start gegaan en kunnen daarom geen data over 2009 aanleveren. En, hoewel sommige instellingen wel hebben deelgenomen aan een pilot CQI-meting, zijn zij niet allen in staat geweest om gegevens aan te leveren, omdat zij niet de beschikking hadden over gecasemixte schaalscores die door de meetorganisatie aangeleverd hadden moeten worden.

Conclusies

De mate waarin prestatie-indicatoren aangeleverd worden kan voor nagenoeg alle indicatoren worden verbeterd. Gemiddeld leveren instellingen gegevens aan voor iets meer dan de helft van alle indicatoren die op de instelling van toepassing zijn. Slechts 10 van de 38 indicatoren is door meer dan 80% van de instellingen aangeleverd. Met name de facultatief aan te leveren prestatie-indicatoren zijn door de meeste instellingen open gelaten. Maar ook een groot aantal verplichte indicatoren is slechts door een klein aantal van de instellingen aangeleverd, in veel gevallen omdat de gegevens niet geregistreerd zijn of omdat de informatie lastig uit de registratie te halen is. Daarnaast is aanlevering op een aantal indicatoren incompleet, omdat deze indicatoren slechts op een beperkt deel van de instellingen van toepassing blijken te zijn.

2.4 Betrouwbaarheid

Programma Zichtbare Zorg heeft de afgelopen jaren een aanpak ontwikkeld om de kwaliteit van de data vast te stellen. De aanpak is nog in ontwikkeling. Dit rapport is een eerste aanzet om de analyses van de prestatie-indicatoren GGZ over het verslagjaar 2009 op een, met andere sectoren vergelijkbare, manier uit te voeren. Een van de elementen waarop de kwaliteit kan worden vastgesteld, is het bepalen van de betrouwbaarheid van de indicatoren.

De betrouwbaarheid van indicatoren kan op verschillende manieren worden gezien. In dit hoofdstuk wordt als eerste ingegaan op de randvoorwaarden voor een statistisch betrouwbare vergelijkbaarheid van instellingen met het overall gemiddelde. Om deze vergelijking verantwoord te kunnen maken, moet er sprake zijn van een voldoende aantal observaties binnen instellingen. De algemene systematiek wordt besproken en per indicator wordt aangegeven in hoeverre aan de voorwaarden voor statistisch betrouwbare vergelijkbaarheid wordt voldaan.

Als tweede wordt in dit hoofdstuk een analyse gemaakt van het onderscheidend vermogen van de indicatoren. Indicatoren zijn immers ontwikkeld om verschillen tussen instellingen inzichtelijk te maken; daarvoor is nodig dat ook daadwerkelijk verschillen tussen instellingen bestaan. Wanneer is vastgesteld dat de indicatoren aan de minimale eisen voor statistisch betrouwbare vergelijkbaarheid voldoen, wordt bekeken in hoeverre instellingen significant verschillen van het gemiddelde van de instellingen samen.

Als derde wordt de 'hertest betrouwbaarheid van indicatoren' bestudeerd. Deze analyse is gedaan voor indicatoren die in zowel 2008 als 2009 op een vergelijkbare manier zijn uitgevraagd. Berekend is in hoeverre instellingen in beide jaren eenzelfde vergelijkbare onderlinge rangorde kennen en in hoeverre instellingen dezelfde indicatorwaarden hebben in beide jaren.

2.4.1 Statistisch betrouwbare vergelijkbaarheid

Aanpak

Bij kwantitatieve onderzoeken wil men vaak toetsen in hoeverre groepen van waarnemingen van elkaar verschillen op vastgestelde parameters. Een voorbeeld van een dergelijke toets is of de indicatorwaarde van een instelling verschilt van een normwaarde. Om een dergelijke uitspraak statistisch verantwoord te doen, is er een minimumaantal waarnemingen per instelling nodig. Om inzichtelijk te maken in hoeverre instellingen voldoen aan het minimumaantal waarnemingen, is bij Zichtbare Zorg Ziekenhuizen de signaalvlag statistisch betrouwbare vergelijkbaarheid ontwikkeld. Deze signaalvlag voor statistisch betrouwbare vergelijkbaarheid wordt nu eveneens gebruikt voor Zichtbare Zorg GGZ.

De vlaggen doen nadrukkelijk geen uitspraak over de kwaliteit van zorg, maar maken mogelijke beperkingen bij de interpretatie van data inzichtelijk. Omdat het proces van transparantie in kwaliteit van zorg nog in ontwikkeling is, zijn de aangeleverde data nog niet altijd voldoende betrouwbaar, valide en vergelijkbaar. Vooral organisaties die de kwaliteitsdata nu al wel willen gebruiken voor bijvoorbeeld de doorvertaling naar publieksinformatie of informatie ten behoeve van zorginkoop, hebben baat bij de signaalvlaggen. De vlaggen bieden namelijk inzicht in de zeggingskracht van de aangeleverde gegevens. In het geval van statistische betrouwbare vergelijkbaarheid van

de indicatorwaarden gaat het erom dat de indicatorwaarde nauwkeurig dient te zijn gemeten en vergelijking met het landelijke gemiddelde statistisch verantwoord moet zijn. De signaalvlag kent drie mogelijke kleuren die elk hun eigen betekenis hebben:

- Rood: alle instellingen bezien, is een zinnige vergelijking van instellingen met het landelijke gemiddelde niet mogelijk.
- Oranje: alle instellingen bezien, vraagt een zinnige vergelijking van instellingen met het landelijke gemiddelde een genuanceerd oordeel.
- Groen: alle instellingen bezien, is een zinnige vergelijking van instellingen met het landelijke gemiddelde op statistisch verantwoord mits de populaties van de instellingen vergelijkbaar zijn.

De kwaliteitsdimensie statistisch betrouwbare vergelijkbaarheid speelt alleen een rol bij indicatoren die als een proportie worden uitgevraagd. Dit kun je interpreteren als een soort van ja/nee-vragen, die leiden tot een teller met bijvoorbeeld het aantal keren 'ja' en een noemer voor het totale aantal patiënten waarop de vraag betrekking heeft.

Het is dus niet altijd mogelijk om signaalvlaggen uit te rekenen. Het gaat om de volgende gevallen:

- Voor indicatoren die niet als proportie worden uitgevraagd, is het niet mogelijk om een vlag voor statistische betrouwbare vergelijkbaarheid te berekenen.
- Voor indicatoren die worden uitgevraagd als een gemiddelde, zonder de bijbehorende spreidingsmaat, is het niet mogelijk een vlag voor statistisch betrouwbare vergelijkbaarheid te berekenen.
- Voor indicatoren die met een eigen meetinstrument zijn gemeten, is een vergelijking op inhoudelijke gronden niet zinvol, omdat het om verschillende indicatoren kan gaan.

De vlag voor statistisch betrouwbare vergelijkbaarheid wordt per instelling als volgt berekend:

- Bij een overall proportie groter dan of gelijk aan 0.05 (5%), maar kleiner of gelijk aan 0.95 (95%):
 - o De statistisch betrouwbare vergelijkbaarheid wordt 'goed' genoemd wanneer het totaal aantal waarnemingen groter of gelijk is aan 200 (dit komt overeen met een Cohen's effect size van 0.2, een klein effect en een odds ratio van ongeveer 2); de instelling krijgt dan een groene signaalvlag voor statistisch betrouwbare vergelijkbaarheid;
 - o De statistisch betrouwbare vergelijkbaarheid wordt 'slecht' genoemd wanneer het totaal aantal waarnemingen kleiner of gelijk is aan 30 (dit komt overeen met een Cohen's effect size van 0.5, een middelmatig effect en een odds ratio van ongeveer 4); de instelling krijgt dan een rode signaalvlag voor statistisch betrouwbare vergelijkbaarheid;
 - o Ligt het aantal waarnemingen tussen de 30 en 200, dan krijgt de instelling een oranje signaalvlag voor statistisch betrouwbare vergelijkbaarheid, de statistisch betrouwbare vergelijkbaarheid is 'matig' te noemen.
- Bij een overall proportie kleiner dan 0.05 (5%) of groter dan 0.95 (95%):
 - o De statistisch betrouwbare vergelijkbaarheid wordt 'goed' genoemd wanneer het totaal aantal waarnemingen groter of gelijk is aan 800 (dit komt overeen met een Cohen's effect size van 0.1, een erg klein effect en een odds ratio van ongeveer 2); de instelling krijgt dan een groene signaalvlag voor statistisch betrouwbare vergelijkbaarheid;
 - o De statistisch betrouwbare vergelijkbaarheid wordt 'slecht' genoemd wanneer het totaal aantal waarnemingen kleiner of gelijk is aan 200 (dit

komt overeen met een Cohen's effect size van 0.2, een klein effect en een odds ratio van ongeveer 4); de instelling krijgt dan een rode signaalvlag voor statistisch betrouwbare vergelijkbaarheid;

- o Ligt het aantal waarnemingen tussen de 200 en 800, dan krijgt de instelling een oranje signaalvlag voor statistisch betrouwbare vergelijkbaarheid, de statistisch betrouwbare vergelijkbaarheid is 'matig' te noemen.

De beoordeling van de statistische betrouwbaarheid voor de indicator - voor alle instellingen

gezamenlijk - wordt gebaseerd op de signaalvlaggen per instelling:

- Wanneer 25% of meer van de instellingen een rode signaalvlag heeft, dan is de statistische betrouwbaarheid voor de betreffende indicator 'slecht' en krijgt deze een rode signaalvlag;
- Wanneer 75% of meer van de instellingen een groene signaalvlag heeft, dan krijgt de indicator een groene signaalvlag en is de statistische betrouwbaarheid 'goed';
- Als de signaalvlag niet rood en niet groen is. Dus: minder dan 25% van de instellingen rood en minder dan 75% van de instellingen groen, dan krijgt de indicator een oranje signaalvlag en is de statistische betrouwbaarheid 'matig'.

Resultaten

Tabellen in een bijlage geven de signaalvlaggen voor statistisch betrouwbare vergelijkbaarheid weer. Steeds is aangegeven of voor de betreffende indicator een signaalvlag voor statistisch betrouwbare vergelijkbaarheid kan worden uitgerekend. Daar waar de vlaggen zijn uitgerekend, wordt de gemiddelde indicatorwaarde, het aantal instellingen (N), het percentage instellingen met een rode, oranje en groene vlag weergegeven. Tot slot wordt in kleur de overall vlag voor de indicator weergegeven.

Conclusies

Opvallend is dat geen van de indicatoren op overall niveau een groene signaalvlag heeft voor statistisch betrouwbare vergelijkbaarheid. De aanleiding hiervoor is dat veel van de aanleverende GGZ-instellingen beperkt in omvang zijn en daardoor beperkte aantallen observaties kennen binnen de instellingen. Het toetsen van verschillen tussen instellingen en het overall gemiddelde voor een indicator moet daarom genuanceerd beoordeeld worden. Die signaalvlag is oranje voor de volgende indicatoren:

- 2a Oordeel cliënt verandering ernst problematiek, voor de aanlevering via de Cliëntthermometer
- 1.2b Systematisch meten verandering ernst problematiek voor de doelgroepen met stemmingsstoornissen, angststoornissen, overige stoornissen en de restgroep
- 1.2d Vermindering van klachten voor de doelgroep met overige stoornissen
- 1.2d Stabilisatie van klachten voor de doelgroep met overige stoornissen
- 1.5a Percentage metingen naar onvervulde zorgbehoefte
- 1.7 Drop-out voor de doelgroep met stoornissen in de kindertijd, stemmingsstoornissen, angststoornissen, overige stoornissen, restgroep
- 1.11 Tijdig contact na ontslag uit kliniek, voor de doelgroepen met schizofrenie, stemmingsstoornissen, aan een middel gebonden stoornis, overige stoornissen
- 2.2 Informatie over bijwerkingen medicijnen (CQI)
- 3.1c Het percentage cliënten waarbij wachttijd tot start behandeling korter is dan x dagen
- 3.3 Informed consent, gemeten met de Cliëntthermometer

- 3.4 Keuzevrijheid, gemeten met de Cliëntthermometer
- 3.5 Vervulling zorgwensen gemeten met de Cliëntthermometer
- 3.6b Percentage geëvalueerde of bijgestelde behandel- of begeleidingsplannen
- 3.9 Cliëntoordeel woon- en leefomstandigheden
- 3.10 Adequate bejegening door hulpverleners
- 3.11 Adequate informatieverstrekking door hulpverleners

Voor de overige indicatoren, met een rode overall vlag voor statistisch betrouwbare vergelijkbaarheid, wordt vergelijking van instellingen met het gemiddelde afgeraden.

2.4.2 Onderscheidend vermogen

Aanpak

Om een indruk te vormen van het onderscheidend vermogen van de indicatoren, is berekend in welke mate instellingen significant afwijken van het gewogen gemiddelde van alle instellingen gezamenlijk. Het gewogen gemiddelde is verkregen door het optellen van alle tellers, en het optellen van alle noemers. Om het onderscheidende vermogen vast te stellen, moet een betrouwbaarheidsinterval rond de indicatorwaarde van de instelling berekend worden. Voor indicatorwaarden die uitgedrukt worden als een proportie kan dit betrouwbaarheidsinterval worden berekend als een functie van de teller, de noemer en de benodigde betrouwbaarheid (hier is de gebruikelijke 95% gehanteerd). Voor het berekenen van het betrouwbaarheidsinterval is in deze analyse de methode van Wilson gehanteerd, aangezien die speciaal geschikt is voor kleine aantallen observaties en extreme proporties. Een instelling wordt beschouwd als significant afwijkend van het gewogen gemiddelde van alle instellingen, wanneer het berekende betrouwbaarheidsinterval van de indicatorwaarde van een instelling niet overlapt met het gewogen gemiddelde van alle instellingen.

Het onderscheidend vermogen van een indicator is uitgedrukt als het percentage instellingen dat significant afwijkt van het gewogen gemiddelde van alle instellingen. Hoe hoger het percentage significant afwijkende instellingen voor een indicator, hoe groter het onderscheidend vermogen van deze indicator is. Op basis van toeval zou men, gezien het gehanteerde 95% betrouwbaarheidsinterval, mogen verwachten dat ongeveer 5% van de instellingen significant afwijkt van het gewogen gemiddelde. In een dergelijk geval is het onderscheidend vermogen dus eigenlijk nihil. Wanneer bijvoorbeeld 50% van de instellingen significant afwijkt van het gemiddelde, is dat tien maal zoveel als op basis van kans verwacht mag worden. Zulke hoge percentages afwijkende instellingen wijzen dus op een substantieel onderscheidend vermogen.

Niet voor alle indicatoren is een berekening van het onderscheidend vermogen mogelijk. Voor indicatoren waar sprake is van een gemiddelde zonder de bijbehorende spreidingsmaat, is het niet mogelijk een betrouwbaarheidsinterval te berekenen. Door het ontbreken van een betrouwbaarheidsinterval is het eveneens niet mogelijk om de significantie van de afwijking van de indicatorwaarde van een instelling van het gemiddelde van alle instellingen te bepalen.

Berekening van onderscheidend vermogen dient ook inhoudelijk zinvol te zijn. Voor indicatoren die gaan over een eigen meetinstrument is een vergelijking van instellingen onderling niet zinvol, omdat het om verschillende meetinstrumenten kan gaan en daardoor onvergelijkbare getallen zouden worden vergeleken.

Met het oog op de statistisch betrouwbare vergelijkbaarheid is het niet voor alle indicatoren zinvol om het onderscheidend vermogen van de indicator te bepalen aan de hand van significante verschillen. Een vergelijking is mogelijk wanneer de signaalvlag voor statistisch betrouwbare vergelijkbaarheid groen is, moet genuanceerd beoordeeld worden bij een oranje signaalvlag, en wordt afgeraden voor indicatoren met een rode signaalvlag. Daarom wordt het onderscheidend vermogen hier alleen gepresenteerd voor indicatoren die een oranje of groene signaalvlag hebben. Zie ook de voorgaande paragraaf voor een toelichting op en een overzicht van de signaalvlaggen voor statistisch betrouwbare vergelijkbaarheid.

Resultaten

In onderstaande tabellen worden, voor de indicatoren waarvoor berekening van het onderscheidend vermogen mogelijk is en op basis van de signaalvlag zinvol wordt geacht, drie waarden weergegeven: het aantal instellingen dat een geldige waarde heeft voor de indicator; hoeveel instellingen daarvan significant afwijken van het gewogen gemiddelde; en tot slot welk percentage van de instellingen significant afwijkt van het gewogen gemiddelde.

Indicator	N	N significant	% significant
1.2a Oordeel cliënt verandering ernst problematiek			
Cliëntthermometer	88	47	53%
1.2b Systematisch meten verandering ernst problematiek			
Stemmingsstoornissen	43	38	88%
Angststoornissen	43	37	86%
Overige stoornis(sen)	28	27	96%
Restgroep	14	12	86%
1.2d Vermindering van klachten			
Overige stoornis(sen)	26	20	77%
1.2d Stabilisatie van klachten			
Overige stoornis(sen)	26	18	69%

Indicator	N	N significant	% significant
1.5a Percentage metingen naar onvervulde zorgbehoefte			
	16	12	75%

Indicator	N	N significant	% significant
1.7 Drop-out			
Stoornissen in de kindertijd	72	42	58%
Stemmingsstoornissen	77	41	53%
Angststoornissen	76	40	53%
Overige stoornis(sen)	63	43	68%
Restgroep	19	11	58%

Indicator	N	N significant	% significant
1.11 Tijdig contact na ontslag uit kliniek			
Schizofrenie e.a.	34	21	62%
Stemmingsstoornissen	36	25	69%
Aan een middel gebonden stoornis	31	20	65%

Overige stoornis(sen)	28	18	64%
-----------------------	----	----	-----

Indicator	N	N significant	% significant
2.2 Informatie over bijwerkingen medicijnen (CQI)	23	4	17%

Indicator	N	N significant	% significant
3.1c Het percentage cliënten waarbij wachttijd tot start behandeling korter is dan x dagen			
korter dan 98 dagen - Ambulant	135	119	88%
korter dan 105 dagen - Klinische zorg	55	40	73%

Indicator	N	N significant	% significant
3.3 Informed consent			
Cliëntthermometer	88	54	61%

Indicator	N	N significant	% significant
3.4 Keuzevrijheid			
Cliëntthermometer	93	47	51%

Indicator	N	N significant	% significant
3.5 Vervulling zorgwensen			
Cliëntthermometer	95	53	56%

Indicator	N	N significant	% significant
3.6b Percentage geëvalueerde of bijgestelde behandel- of begeleidingsplannen	99	84	85%

Indicator	N	N significant	% significant
3.9 Cliëntoordeel woon- en leefomstandigheden			
Cliëntthermometer	30	5	17%

Indicator	N	N significant	% significant
3.10 Adequate bejegening door hulpverleners			
Cliëntthermometer	95	45	47%

Indicator	N	N significant	% significant
3.11 Adequate informatieverstrekking door hulpverlener			
Cliëntthermometer	97	60	62%

Conclusies

Voor de meeste indicatoren waarvan het onderscheidend vermogen is bepaald, geldt dat het onderscheidende vermogen rond of boven de 50% zit, wat als substantieel onderscheidend vermogen beschouwd kan worden. Twee indicatoren vormen hierop een uitzondering:

- indicator 2.2 Informatie over bijwerkingen medicijnen, gemeten met de CQ-index;
- indicator 3.9 Cliëntoordeel woon- en leefomstandigheden (Cliëntthermometer).

Voor beide indicatoren geldt dat 17% van de instellingen significant afwijkt van het gemiddelde. Dit is weliswaar meer dan de 5% die op basis van toeval verwacht mag worden, maar beduidend lager dan bij de andere indicatoren is gevonden. Blijkbaar verschillen GGZ-instellingen op deze indicatoren niet vaak significant van elkaar.

2.4.3 Correlaties tussen de meetjaren

Aanpak

Wanneer bij dezelfde instelling in opeenvolgende jaren dezelfde indicator wordt uitgevraagd, mag men verwachten dat de samenhang tussen de indicatorwaarden van de verschillende meetjaren hoog is. De mate van samenhang is daarmee een schatting van de betrouwbaarheid waarmee een indicator kan worden gemeten. Is de samenhang relatief laag, dan kan dit twee dingen betekenen:

- Het kan zijn dat er grote veranderingen zijn geweest binnen de instellingen in de onderzochte periode, die verklaren dat de observaties tussen de jaren verschillen.
- Het kan zijn dat de veranderingen binnen de instellingen beperkt zijn. Dit is vermoedelijk doorgaans het geval, aangezien de voorgaande meting een jaar geleden is uitgevoerd. Dan is bij een lage samenhang kennelijk de gemeten indicatorwaarde sterk van het toeval afhankelijk en daardoor niet goed betrouwbaar te meten.

Voor het bepalen van de samenhang worden in dit onderzoek twee verschillende maten gebruikt, die ieder een eigen betekenis hebben: de Spearman-rangcorrelatie, en de intraklassecorrelatie (ICC).

De Spearman-correlatie is een rangcorrelatiemaat en geeft een antwoord op de vraag in hoeverre instellingen in beide meetjaren in dezelfde volgorde staan wanneer zij gesorteerd worden op de indicatorwaarde. Een correlatie van 1 duidt op een perfecte positieve samenhang (alle instellingen staan in beide jaren in exact dezelfde volgorde) en een correlatie van -1 duidt op een perfecte negatieve samenhang (de instellingen staan in exact de omgekeerde volgorde). Wanneer de rangorde van de twee meetmomenten in het geheel niet samenhangen, is de correlatie 0.

De Spearman-correlatie gaat over relatieve samenhang in de zin van rangordening, niet over de mate waarin instellingen in beide meetjaren identieke scores hebben in absolute zin. De ICC drukt uit in hoeverre de variatie tussen de observaties in beide meetjaren zijn toe te schrijven aan de instellingen in termen van een proportie van de variatie. Een ICC van 1 geeft aan dat instellingen in beide meetjaren exact dezelfde waarden hebben en alle verschillen in observaties liggen tussen de instellingen, terwijl een ICC van 0 impliceert dat geen verschillen zijn toe te schrijven aan instellingen.

De indicatoren die over 2008 en 2009 zijn uitgevraagd zijn voor een deel inhoudelijk goed vergelijkbaar, daarvoor zijn deze correlaties uitgerekend. Voor een ander deel zijn de indicatoren inhoudelijk dusdanig veranderd dat een vergelijking niet mogelijk is en correlaties geen inhoudelijke betekenis hebben. De volgende tabel geeft weer voor welke indicatoren een vergelijking tussen de meetjaren op inhoudelijke gronden mogelijk is.

Indicator	Omschrijving
Deel 1	Effectiviteit
1.2a	Oordeel van de cliënt over verandering ernst problematiek
1.2b	Systematisch meten verandering ernst problematiek
1.2c	Gebruikte instrumenten
1.2d	Verandering ernst problematiek
1.4b	Gebruikte instrumenten
1.7	Drop-out
1.11	Tijdig contact na ontslag uit kliniek
Deel 2	Veiligheid
2.2	Informatie over bijwerkingen medicijnen
2.3	Onveiligheid door ontbreken van informatie
Deel 3	Cliëntgerichtheid
3.1c	Percentage wachttijd korter dan x dagen
3.3	Informed consent
3.4	Keuzevrijheid
3.5	Vervulling zorgwensen
3.9	Cliëntoordeel woon- en leefomstandigheden in klinische of RIBW-setting
3.10	Adequate bejegening door de hulpverlener
3.11	Adequate informatieverstrekking over de behandeling/begeleiding

Om een vergelijking van deze indicatoren tussen de meetjaren 2008 en 2009 mogelijk te maken, zijn de resultaten van de instellingen in de datasets van beide jaren aan elkaar gekoppeld op basis van de KvK-nummers en de namen van de instellingen. Gefuseerde instellingen zijn omwille van de vergelijkbaarheid buiten beschouwing gelaten. In de onderstaande tabellen is per indicator steeds het aantal gekoppelde instellingen dat in beide jaren de indicator heeft aangeleverd (N), de Spearman-correlatie en de bijbehorende significantie (p), en de ICC uitgerekend. Aangezien zowel de Spearman-correlatie als de ICC slechts zinvol zijn wanneer er sprake is van een enigszins substantieel aantal observaties, zijn de berekende schattingen alleen weergegeven wanneer er minimaal 20 observaties zijn. Wanneer voor een (deel)indicator geen vergelijking mogelijk bleek, is dit eveneens vermeld.

Resultaten verkregen met een eigen meetinstrument zijn in deze analyse wel meegenomen. Aangenomen wordt dat de vraagstelling binnen een instelling in beide meetmomenten gelijk is gebleven.

Resultaten

Onderstaande tabellen geven de resultaten van de berekende Spearman-correlaties en de ICC.

Indicator	N	Spearman	p	ICC
1.2a Oordeel cliënt verandering ernst problematiek				
Cliëntthermometer	51	0,72	0,00	0,72
Eigen meetinstrument	17	aantal observaties te laag		
CQ-index		geen vergelijking mogelijk		
1.2b Systematisch meten verandering ernst problematiek				
Stoornissen in de kindertijd	12	aantal observaties te laag		
Schizofrenie e.a.	9	aantal observaties te laag		
Stemmingsstoornissen	21	0,63	0,00	0,54

Indicator	N	Spearman	p	ICC
Angststoornissen	19	aantal observaties te laag		
Persoonlijkheidsstoornissen	14	aantal observaties te laag		
Delirium, dementie e.a.	8	aantal observaties te laag		
Aan een middel gebonden stoornis	8	aantal observaties te laag		
Overige stoornis(sen)	9	aantal observaties te laag		
Restgroep	2	aantal observaties te laag		
1.2c Gebruikte instrumenten				
Dichotome indicator, geen correlatie mogelijk				
1.2d Vermindering van klachten				
Stoornissen in de kindertijd	4	aantal observaties te laag		
Schizofrenie e.a.	5	aantal observaties te laag		
Stemmingsstoornissen	13	aantal observaties te laag		
Angststoornissen	12	aantal observaties te laag		
Persoonlijkheidsstoornissen	7	aantal observaties te laag		
Delirium, dementie e.a.	3	aantal observaties te laag		
Aan een middel gebonden stoornis	4	aantal observaties te laag		
Overige stoornis(sen)	6	aantal observaties te laag		
Restgroep	1	aantal observaties te laag		
1.2d Stabilisatie van klachten				
Stoornissen in de kindertijd	4	aantal observaties te laag		
Schizofrenie e.a.	5	aantal observaties te laag		
Stemmingsstoornissen	13	aantal observaties te laag		
Angststoornissen	12	aantal observaties te laag		
Persoonlijkheidsstoornissen	8	aantal observaties te laag		
Delirium, dementie e.a.	3	aantal observaties te laag		
Aan een middel gebonden stoornis	4	aantal observaties te laag		
Overige stoornis(sen)	6	aantal observaties te laag		
Restgroep	1	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
1.4b Gebruikte instrumenten				
Dichotome indicator, geen correlatie mogelijk				

Indicator	N	Spearman	p	ICC
1.7 Drop-out				
Stoornissen in de kindertijd	44	0,65	0,00	0,55
Schizofrenie e.a.	34	0,33	0,05	0,13
Stemmingsstoornissen	47	0,79	0,00	0,76
Angststoornissen	46	0,80	0,00	0,93
Persoonlijkheidsstoornissen	37	0,67	0,00	0,60
Delirium, dementie e.a.	32	0,68	0,00	0,26
Aan een middel gebonden stoornis	34	0,74	0,00	0,78
Overige stoornis(sen)	32	0,67	0,00	0,59
Restgroep	3	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
1.11 Tijdig contact na ontslag uit kliniek				
Stoornissen in de kindertijd	26	0,48	0,01	0,19
Schizofrenie e.a.	25	0,60	0,00	0,57
Stemmingsstoornissen	26	0,51	0,01	0,42
Angststoornissen	27	0,19	0,34	-
Persoonlijkheidsstoornissen	20	0,35	0,12	0,53
Delirium, dementie e.a.	22	0,39	0,07	0,32
Aan een middel gebonden stoornis	19	aantal observaties te laag		
Overige stoornis(sen)	15	aantal observaties te laag		
Restgroep	1	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
2.2 Informatie over bijwerkingen medicijnen (CQI)	5	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
2.3 Onveiligheid door ontbreken van informatie	11	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
3.1c Het percentage cliënten waarbij wachttijd tot start behandeling korter is dan x dagen				
korter dan 98 dagen - Ambulant	75	0,41	0,00	0,25
korter dan 105 dagen - Klinische zorg	38	0,62	0,00	0,61
korter dan 147 dagen - RIBW	13	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
3.3 Informed consent				
Cliëntthermometer	55	0,68	0,00	0,74
Eigen meetinstrument	13	aantal observaties te laag		
CQ-index	5	aantal observaties te laag		

Indicator	N	Spearman	p	ICC
3.4 Keuzevrijheid				
Cliëntthermometer	59	0,73	0,00	0,75
Eigen meetinstrument	16	aantal observaties te laag		
CQ-index		geen vergelijking mogelijk		

Indicator	N	Spearman	p	ICC
3.5 Vervulling zorgwensen				
Cliëntthermometer	60	0,79	0,00	0,77
Eigen meetinstrument	19	aantal observaties te laag		
CQ-index		geen vergelijking mogelijk		

Indicator	N	Spearman	p	ICC
3.9 Cliëntoordeel woon- en leefomstandigheden				

Cliëntthermometer	18	aantal observaties te laag
Eigen meetinstrument	6	aantal observaties te laag
CQ-index		geen vergelijking mogelijk

Indicator	N	Spearman	p	ICC
3.10 Adequate bejegening door hulpverleners				
Cliëntthermometer	59	0,76	0,00	0,58
Eigen meetinstrument	21	0,48	0,03	-
CQ-index		geen vergelijking mogelijk		

Indicator	N	Spearman	p	ICC
3.11 Adequate informatieverstrekking door hulpverlener				
Cliëntthermometer	60	0,71	0,00	0,65
Eigen meetinstrument	20	0,36	0,12	-
CQ-index		geen vergelijking mogelijk		

Conclusies

In 40 gevallen blijkt het aantal instellingen dat in beide meetjaren de betreffende indicator heeft aangeleverd lager dan 20. Daarnaast is voor zes indicatoren die afgeleid zijn van de CQ-index geen vergelijking mogelijk, omdat er geen corresponderende variabelen zijn gevonden in beide meetjaren. Voor dichotome indicatoren is eveneens geen correlatie berekend. Van de overige 26 indicatoren is de samenhang tussen beide meetjaren berekend. Voor twee van deze indicatoren bleek het niet mogelijk een ICC te berekenen.

De Spearman-rangcorrelaties zijn over het algemeen significant positief ($p < .05$). Dit betekent dat, zoals verwacht, de gevonden rangorden van instellingen in beide meetjaren enigszins vergelijkbaar zijn. Ook de berekende ICC's zijn over het algemeen substantieel. Hieruit kan men concluderen dat ook de indicatorwaarden binnen een instelling in de opeenvolgende jaren relatief sterk op elkaar lijken. Voor de meeste van deze indicatoren geldt zodoende dat de samenhang tussen de meetjaren wijst op een redelijke tot goede betrouwbaarheid.

Waar precies de grens moet worden getrokken voor een goede en een minder goede betrouwbaarheid blijft enigszins arbitrair. Wanneer zoals hier dezelfde indicatoren meermalen bij dezelfde instellingen worden uitgevraagd, mag men echter een positieve Spearman-correlatie verwachten van boven de 0,50. Voor acht indicatoren wordt een Spearman-rangcorrelatie gevonden van lager dan 0,50. Een ICC lager dan 0,50 betekent dat minder dan de helft van de verschillen tussen de observaties kan worden toegeschreven aan de instellingen, en dus meer dan de helft van de verschillen aan andere oorzaken dient te worden toegeschreven. Voor zes indicatoren is de ICC lager dan 0,50.

De indicatoren die een Spearman-correlatie lager dan 0,50 en / of een ICC lager dan 0,50 hebben zijn:

- indicator 1.7 Drop-out (diagnosegroepen schizofrenie e.a., evenals delirium, dementie e.a.);

- indicator 1.11 Tijdig ontslag uit de kliniek (diagnosegroepen stoornissen in de kindertijd, stemmingsstoornissen, angststoornissen, persoonlijkheidsstoornissen, evenals delirium, dementie, e.a.);
- indicator 3.1c Het percentage cliënten waarbij wachttijd tot start behandeling korter is dan x dagen (korter dan 98 dagen - Ambulant);
- indicator 3.10 Adequate bejegening door hulpverleners (eigen meetinstrument);
- indicator 3.11 Adequate informatieverstrekking door hulpverlener (eigen meetinstrument).

Het verdient aanbeveling om de bruikbaarheid van de deelindicatoren die niet voldoen aan de hier gehanteerde grenzen voor minimale betrouwbaarheid opnieuw in overweging te nemen. Een lage Spearman-rangcorrelatie wijst erop dat de rangorde van jaar tot jaar meer verschilt dan men zou mogen verwachten. Een lage ICC wijst erop dat herhaalde meting binnen dezelfde instellingen grotere verschillen oplevert tussen meetmomenten dan tussen de instellingen.

2.4 Validiteit en populatievergelijkbaarheid

Aanpak

In dit hoofdstuk worden de indicatoren aan de hand van de deelaspecten validiteit en populatievergelijkbaarheid bekeken. Deze twee deelaspecten zijn in evaluatiebijeenkomsten met, veelal kwaliteitsmedewerkers van, GGZ-instellingen vastgesteld op het niveau van de indicator. Om een completer beeld te krijgen van de validiteit gaat Zichtbare Zorg in 2011 de validiteit met een bredere groep vaststellen, waarin naast kwaliteitsmedewerkers van GGZ-instellingen, ook de input van het cliëntperspectief wordt meegenomen.

Er dient opgemerkt te worden dat de projectgroep in het najaar van 2010 heeft bepaald welke indicatoren er definitief vervallen en welke indicatoren er verder worden ontwikkeld. De input vanuit de evaluatiebijeenkomsten zijn meegenomen in de bepaling.

Validiteit zegt iets over de mate waarin een indicator (geoperationaliseerde grootheid) meet wat hij moet meten (abstract concept). Uitspraak hierover is gebaseerd op de input die GGZ-instellingen hebben geleverd in de evaluatiebijeenkomsten. Om de validiteit te bepalen, is voor iedere indicator vastgesteld of er bij de GGZ-instellingen duidelijkheid was over:

1. De interpretatie van de indicator;
2. De in- en exclusiecriteria die toegepast moesten worden;
3. De relatie tussen de indicator en de kwaliteit van zorg.

De door de GGZ-instellingen aangegeven relatie met de kwaliteit van zorg, is het belangrijkste onderdeel. Als er voor een indicator géén relatie is met de kwaliteit van zorg, dan dient vanuit het perspectief van de validiteit de indicator grondig te worden herzien. Als er wél een relatie is met de kwaliteit van zorg, dan geven de duidelijkheid over de interpretatie en de in- en exclusiecriteria richting voor het verbeteren van de bestaande indicator.

Populatievergelijkbaarheid zegt iets over de mate waarin factoren, die kunnen verschillen tussen GGZ-instellingen (bijvoorbeeld patiëntenkenmerken), van invloed zijn op de indicatorwaarde. In de evaluatiebijeenkomsten zijn door de GGZ-instellingen meerdere malen opmerkingen gemaakt over de populatievergelijkbaarheid. Hoewel het onderwerp van populatievergelijkbaarheid niet expliciet is besproken in de evaluatiebijeenkomsten, leggen we in deze rapportage een eerste aanzet tot vaststelling van de populatievergelijkbaarheid vast.

Om de validiteit van de indicatoren vast te stellen, is een drietal evaluatiebijeenkomsten gehouden met GGZ-instellingen. De bijeenkomsten zijn in augustus 2010 gehouden. GGZ-instellingen vaardigden veelal kwaliteitsmedewerkers af om de bijeenkomsten bij te wonen. Om een representatieve vertegenwoordiging te krijgen van de instellingen, is op basis van een steekproef een aantal instellingen uitgenodigd deel te nemen aan de bijeenkomsten. Bij het trekken van de steekproef is gekeken naar de volgende kenmerken van de instellingen: omvang, aantal scholingsvragen dat is gesteld, de mate waarin de instellingen de pagina's op de webenquête hebben ingevuld en of er forensische zorg wordt geleverd. Daarbuiten konden overige GGZ-instellingen zich ook voor de bijeenkomsten aanmelden. Instellingen die niet aanwezig konden zijn, hebben via e-mail input kunnen aanleveren. Van de instellingen waren veelal kwaliteitsmedewerkers aanwezig.

Het doel van de evaluatiebijeenkomsten was tweeledig. Ten eerste werd de validiteit van de indicatoren vastgesteld. Het resultaat daarvan is opgenomen in het voor u liggende rapport. Ten tweede werden de verbeterpunten van de indicatoren geïnventariseerd. De verbeterpunten heeft Zichtbare Zorg verwerkt in de indicatorset zoals deze in 2011 over het verslagjaar 2010 zal worden uitgevraagd.

Resultaten

Onderstaande tabel toont de resultaten wat betreft de validiteit van de indicatoren, zoals vastgesteld in de evaluatiebijeenkomsten met GGZ-instellingen.

Indicator	Omschrijving	Interpretatie duidelijk	In- en exclusie criteria duidelijk	Relatie met kwaliteit van de zorg
Deel 1	Effectiviteit			
1.1	Bereik preventieactiviteiten	nee	nee	nee
1.2a	Oordeel van de cliënt over verandering ernst problematiek	ja	nee	ja
1.2b	Systematisch meten verandering ernst problematiek	ja	ja	ja
1.2c	Gebruikte instrumenten	ja	ja	ja
1.2d	Verandering ernst problematiek	ja	ja	nee
1.3	Verandering in het dagelijks functioneren van de cliënt	ja	ja	nee
1.4a	Systematisch meten van verandering in ervaren kwaliteit	ja	ja	ja
1.4b	Gebruikte instrumenten	ja	ja	ja
1.4c	Verandering in de ervaren kwaliteit van leven van de cliënt van leven	ja	ja	ja
1.5a	Percentage metingen naar onvervulde zorgbehoefte in kader rehabilitatie	nee	ja	nee
1.5b	Gebruikte instrumenten	ja	ja	ja
1.6	Heropname versus opname-duur (in ontwikkeling)	Niet uitgevraagd 2009		
1.7	Drop-out	ja	nee	nee
1.8a	De beschikbaarheid van een procedure voor screening op somatische aandoeningen	ja	ja	nee
1.8b	De mate waarin screening op somatische aandoeningen plaatsvindt	ja	ja	ja
1.9	Hanteren behandelrichtlijnen	Niet uitgevraagd 2009		
1.10	Bereik zorgwekkende zorgmijders (in ontwikkeling)	Niet uitgevraagd 2009		
1.11	Tijdig contact na ontslag uit kliniek	ja	ja	ja
Deel 2	Veiligheid			
2.1a	Beschikbaarheid medicatie-overzicht	ja	ja	nee
2.1b	Combinatie SSRI/NSAID met PPI	ja	ja	nee
2.1c	Rationeel voorschrijven benzodiazepines	ja	ja	nee
2.2	Informatie over bijwerkingen medicijnen	ja	ja	ja
2.3	Onveiligheid door ontbreken van informatie	ja	ja	nee
2.4a	Aantal insluitingen/toepassing dwangmedicatie	nee	ja	ja
2.4b	Gemiddelde duur insluiting/dwangmedicatie	ja	ja	ja
2.4c	Evaluatie met cliënt	ja	ja	ja
2.4d	Aanwezigheid preventieprogramma	ja	ja	nee
2.5	Suicidepreventie (in ontwikkeling)	Niet uitgevraagd 2009		
2.6a	Ervaren veiligheid cliënten	ja	ja	ja
2.6b	Perspectief hulpverlener over veiligheid (in ontwikkeling)	Niet uitgevraagd 2009		
Deel 3	Cliëntgerichtheid			
3.1a	Gemiddelde aanmeldingswachttijd	nee	nee	ja
3.1b	Gemiddelde wachttijd tot behandeling	ja	nee	ja
3.1c	Percentage wachttijd korter dan x dagen	ja	nee	ja
3.2	Toegang tot zorg (in ontwikkeling)	Niet uitgevraagd 2009		
3.3	Informed consent	ja	nee	ja
3.4	Keuzevrijheid	ja	ja	ja
3.5	Vervulling zorgwensen	ja	ja	ja
3.6a	Cliëntoordeel evaluatie/bijstelling behandel-/begeleidingsplan	ja	ja	ja
3.6b	Percentage geëvalueerde/bijgestelde behandel-/begeleidingsplannen	ja	ja	ja
3.7a	Ambulante zorg	ja	ja	ja
3.7b	Langdurige zorg/klinische zorg	ja	nee	ja
3.8a	Cliëntervaring deelname aan dag- of werkactiviteiten	ja	ja	ja
3.8b	Perspectief hulpverlener cliëntparticipatie aan dag- of werkactiviteiten (in ontwikkeling)	Niet uitgevraagd 2009		
3.9	Cliëntoordeel woon- en leefomstandigheden in klinische of RIBW-setting	ja	ja	ja
3.10	Adequate bejegening door de hulpverlener	ja	ja	ja
3.11	Adequate informatieverstrekking over de behandeling/begeleiding	ja	nee	ja

In onderstaande tabel zijn de eerste aanwijzingen op het gebied van de populatievergelijkbaarheid samengevat.

Indicator	Omschrijving	Opmerking over populatie vergelijkbaarheid
Deel 1	Effectiviteit	
1.1	Bereik preventieactiviteiten	
1.2a	Oordeel van de cliënt over verandering ernst problematiek	
1.2b	Systematisch meten verandering ernst problematiek	maak onderscheid tussen kort- en langdurige zorg
1.2c	Gebruikte instrumenten	
1.2d	Verandering ernst problematiek	maak onderscheid tussen kort- en langdurige zorg
1.3	Verandering in het dagelijks functioneren van de cliënt	maak onderscheid naar zorgsetting
1.4a	Systematisch meten van verandering in ervaren kwaliteit	
1.4b	Gebruikte instrumenten	
1.4c	Verandering in de ervaren kwaliteit van leven van de cliënt van leven	maak onderscheid naar zorgzwaarte
1.5a	Percentage metingen naar onvervulde zorgbehoefte in kader rehabilitatie	
1.5b	Gebruikte instrumenten	
1.6	Heropname versus opname-duur (in ontwikkeling)	Niet uitgevraagd 2009
1.7	Drop-out	patientenmix is belangrijk
1.8a	De beschikbaarheid van een procedure voor screening op somatische aandoeningen	
1.8b	De mate waarin screening op somatische aandoeningen plaatsvindt	
1.9	Hanteren behandelrichtlijnen	Niet uitgevraagd 2009
1.10	Bereik zorgwekkende zorgmijders (in ontwikkeling)	Niet uitgevraagd 2009
1.11	Tijdig contact na ontslag uit kliniek	deze indicator is vooral belangrijk voor suicidale cliënten
Deel 2	Veiligheid	
2.1a	Beschikbaarheid medicatie-overzicht	
2.1b	Combinatie SSRI/NSAID met PPI	
2.1c	Rationeel voorschrijven benzodiazepines	
2.2	Informatie over bijwerkingen medicijnen	
2.3	Onveiligheid door ontbreken van informatie	
2.4a	Aantal insluitingen/toepassing dwangmedicatie	trend is belangrijker dan benchmark
2.4b	Gemiddelde duur insluiting/dwangmedicatie	
2.4c	Evaluatie met cliënt	
2.4d	Aanwezigheid preventieprogramma	
2.5	Suicidepreventie (in ontwikkeling)	Niet uitgevraagd 2009
2.6a	Ervaren veiligheid cliënten	
2.6b	Perspectief hulpverlener over veiligheid (in ontwikkeling)	Niet uitgevraagd 2009
Deel 3	Cliëntgerichtheid	
3.1a	Gemiddelde aanmeldingswachtijd	
3.1b	Gemiddelde wachttijd tot behandeling	maak onderscheid naar intra- en extramurale zorg
3.1c	Percentage wachttijd korter dan x dagen	maak onderscheid naar regio
3.2	Toegang tot zorg (in ontwikkeling)	Niet uitgevraagd 2009
3.3	Informed consent	
3.4	Keuzevrijheid	
3.5	Vervulling zorgwensen	
3.6a	Cliëntoordeel evaluatie/bijstelling behandel-/begeleidingsplan	
3.6b	Percentage geëvalueerde/bijgestelde behandel-/begeleidingsplannen	
3.7a	Ambulante zorg	
3.7b	Langdurige zorg/klinische zorg	
3.8a	Cliëntervaring deelname aan dag- of werkactiviteiten	
3.8b	Perspectief hulpverlener cliëntparticipatie aan dag- of werkactiviteiten (in ontwikkeling)	Niet uitgevraagd 2009
3.9	Cliëntoordeel woon- en leefomstandigheden in klinische of RIBW-setting	
3.10	Adequate bejegening door de hulpverlener	
3.11	Adequate informatieverstrekking over de behandeling/begeleiding	

Conclusies

Wat betreft de validiteit van de indicatoren zijn er, op basis van de GGZ-instellingen die input hebben geleverd in de evaluatiebijeenkomsten, de volgende conclusies te trekken:

1. Op basis van de vraagtekens bij de relatie van de indicator met de door de instelling beïnvloedbare zorg, verdient het de aanbeveling om de volgende elf indicatoren grondig te herzien:
 - 1.1 Preventieactiviteiten
 - 1.2d Verandering ernst problematiek
 - 1.3 Verandering in het dagelijks functioneren van de cliënt
 - 1.5a Percentage metingen naar onvervulde zorgbehoefte in kader rehabilitatie
 - 1.7 Drop-out
 - 1.8a De beschikbaarheid van een procedure voor screening op somatische aandoening
 - 2.1a Beschikbaarheid medicatie overzicht
 - 2.1b Combinatie SSRI/NSAID met PPI
 - 2.1c Rationeel voorschrijven benzodiazepines
 - 2.3 Onveiligheid door ontbreken van informatie
 - 2.4d Aanwezigheid preventieprogramma
2. Kleinere tekstuele aanpassingen van de indicatoren zijn wenselijk voor de volgende negen indicatoren:
 - 1.2a Oordeel van de cliënt over de verandering ernst problematiek
 - 1.5a Percentage metingen naar onvervulde zorgbehoefte in kader rehabilitatie
 - 2.4a Aantal insluitingen/toepassing dwangmedicatie
 - 3.1a Gemiddelde aanmeldingswachtijd
 - 3.1b Gemiddelde wachttijd tot behandeling
 - 3.1c Percentage wachttijd korter dan x dagen
 - 3.3 Informed consent
 - 3.7b Langdurige zorg/klinische zorg
 - 3.11 Adequate informatieverstrekking over de behandeling/begeleiding
3. Er zijn vraagtekens te zetten bij de populatievergelijkbaarheid van de volgende negen indicatoren:
 - 1.2b Systematisch meten verandering ernst problematiek
 - 1.2d Verandering ernst problematiek
 - 1.3 Verandering in het dagelijks functioneren van de cliënt
 - 1.4c Verandering in de kwaliteit van leven van de cliënt
 - 1.7 Drop-out
 - 1.11 Tijdig contact na ontslag kliniek
 - 2.4a Aantal insluitingen/toepassing dwangmedicatie
 - 3.1b Gemiddelde wachttijd tot behandeling
 - 3.1c Percentage wachttijd korter dan x dagen

2.5 Conclusie

De resultaten van alle uitgevoerde analyses zijn in een grote tabel naast elkaar gelegd. Zie hiervoor Bijlage J. Door het naast elkaar leggen van de resultaten op gebied van volledigheid, betrouwbaarheid en validiteit is het mogelijk om vast te stellen wat het algemene beeld is van de kwaliteit van de indicator. Het spreekt voor zich dat betere scores op alle onderdelen van de gemeten kwaliteit, een positiever beeld van de indicator opleveren. Bij voorkeur zijn de indicatoren waar een positief beeld van is, opgenomen in de kernset van volgend jaar.

Gemiddeld hebben instellingen 61% van de verplichte indicatoren aangeleverd. Een aantal indicatoren is opvallend minder volledig aangeleverd. Het gaat om indicatoren die niet verplicht waren, maar ook wel verplichte indicatoren worden soms minder volledig dan gemiddeld aangeleverd. Dit laatste is opvallend: bij verplichte indicatoren bestaat de verwachting dat deze volledig zouden zijn aangeleverd. Er zijn ook overigens enkele niet verplichte indicatoren die wel meer dan gemiddeld volledig zijn aangeleverd.

Het is opvallend dat de berekende vlaggen op landelijk niveau voor de statistische betrouwbaarheid altijd oranje of rood zijn. Dit betekent dat het toetsen van verschillen tussen instelling genuanceerd moet worden beoordeeld of wordt afgeraden.

De relatie met de kwaliteit van zorg is voor de GGZ instellingen onduidelijk voor een aantal indicatoren uit de onderdelen 1 Effectiviteit en 2 Veiligheid van de Basisset. Voor alle indicatoren uit onderdeel 3 (Cliëntgerichtheid) is die relatie wel duidelijk.

Er zijn geen indicatoren die op alle kwaliteitsaspecten goed scoren.

3 Aanbevelingen

3.1 Procesmatige verbeterpunten

1. Momenteel is er geen zicht op of a) Alle instellingen die verplicht zijn de Basisset Prestatie Indicatoren GGZ aan te leveren, ook daadwerkelijk aanleveren, en b) Alle instellingen die de set hebben aangeleverd, daadwerkelijk de set moesten aanleveren. Het is wenselijk om daar wel zicht op te krijgen.
2. Instellingen kenmerken zichzelf via tien introductievragen. Een deel van de introductievragen is direct gekoppeld aan het van toepassing zijn van indicatoren. Er is geen enkele controle op de juistheid van de aangeleverde kenmerken. Het verdient de aanbeveling om meer grip te krijgen op de kenmerken van de aanleverende instellingen.
3. Stel op verzoek van de GGZ-instellingen via de website van Zichtbare Zorg GGZ een kennisforum in, waar gebruikers elkaar kunnen vinden en praktische vragen kunnen stellen. Dit stelt Zichtbare Zorg tevens in staat om gedurende het gehele proces feeling te houden met het veld.
4. Het verschil in omvang van de instellingen is erg groot. De omvang varieert van zes cliënten per jaar, tot 150.000 cliënten per jaar. Probeer daar meer vergelijkbare eenheden van te maken.
5. Sluit bij het doorontwikkeling van de indicatoren zoveel mogelijk aan op huidige registratiesystemen van instellingen, zodat het voor de instellingen eenvoudig is om de gegevens uit het systeem te halen en aan te leveren.
6. Neem bij de doorontwikkeling van indicatoren in ogenschouw dat aanpassingen invloed hebben op de vergelijkbaarheid met voorgaande jaren en weeg dat af tegen verbetering van de indicator. Juist omdat het veld aangeeft dat in sommige gevallen een van-jaar-tot-jaar vergelijk belangrijker wordt geacht, dan een benchmark van resultaten.

3.2 Inhoudelijke verbeterpunten

1. Instellingen verschillen van elkaar. Dit vormt de aanleiding van het verzoek uit het veld om juist die verschillen inzichtelijk te maken. De manier waarop de indicatoren momenteel worden uitgevraagd, laat dat echter niet toe. Daarbij dient opgemerkt te worden, dat als indicatoren bijvoorbeeld per type instelling uitgevraagd worden, zodat ze ook per type instelling gerapporteerd kunnen worden, de kans reëel is dat er per indicatoren zeer weinig waarnemingen zullen zijn. Bij de doorontwikkeling van de indicatoren is het raadzaam na te gaan of de indicatoren zo kunnen gedefinieerd en op een juist aggregatieniveau worden uitgevraagd, zodat ze recht te kunnen doen aan verschillen tussen de instellingen en dat er voldoende waarnemingen per indicator zijn.
2. De CQ-index data wordt aangeleverd op basis van tellers en noemers, of gemiddelden en noemers. Doordat dit op verschillende manieren wordt uitgevraagd, zijn niet dezelfde analyses toepasbaar op deze vragen. Bovendien heeft het de voorkeur om data op individueel niveau aangeleverd te krijgen, omdat dan betekenisvolle analyses mogelijk zijn. Andere mogelijkheid is het aanleveren van tellers, noemers, aantal cliënten en een spreidingsmaat.
3. De urgentie van aanleveren ontbreekt bij facultatieve indicatoren. Dit beïnvloedt de betrouwbaarheid van de indicator negatief. Het verdient de aanbeveling om deze indicatoren geheel te laten vervallen, en eventueel door te ontwikkelen tot verplichte indicatoren.

4. Gebruik eenduidige, consistentie termen en structuur om de indicatoren te duiden. Stel één gids op met indicatoren. Integreer hiervoor de indicatorgids met de werkinstructie.
5. De noodzaak van het uitvragen van resultaten van een meting met het 'Eigen meetinstrument' kan ter discussie worden gesteld. Resultaten zijn namelijk niet onderling, of met andere resultaten vergelijkbaar.
6. Leg in de indicatorgids van volgend jaar direct vast welke indicatoren op inhoudelijke gronden wel of niet met indicatoren uit voorgaande jaren kunnen worden vergeleken.
7. Voeg vragen toe die een beeld geven van de kwaliteit van de uitvraag analoog aan de verificatievragen die bij de ziekenhuizen worden gebruikt.

3.3 Verbeterpunten ten aanzien van de portal en dataset

1. Maak het mogelijk om bij elke subindicator of elke diagnosegroep opnieuw aan te geven dat er geen gegevens beschikbaar zijn. Zoals het afgelopen jaar is uitgevraagd, moesten instellingen soms waarden aanleveren, terwijl ze dat niet konden. Hierdoor zijn vermoedelijk veel (niet te bepalen hoeveel) nullen ingevoerd.
2. Voor het vaststellen van de volledigheid is het zeer belangrijk dat de routing in de webportal juist is en dat duidelijk wordt aangegeven, zowel op de webportal als in het databestand, als een indicator niet van toepassing is op een instelling.
3. Koppel de verschillende meetjaren van de instellingen in de database aan elkaar, zodat deze koppeling niet achteraf en handmatig gedaan moet worden.
4. Indien de indicator op de instelling van toepassing is en de instelling heeft niet aangegeven dat er geen gegevens beschikbaar zijn, maar de instelling vult toch geen waarden in, laat dan het invulveld leeg. (In SPSS wordt het dan dus 'system missing'). Dit zou dan in principe alleen kunnen voorkomen bij facultatieve indicatoren. Als een indicator verplicht is aan te leveren, zou er altijd een getal moeten staan.
5. Als een instelling na invullen van (een deel van) de indicatorset alsnog het antwoord op een algemene vraag wijzigt, zorg dan dat alle velden die hierdoor niet meer van toepassing zijn op de instelling ook geen waarde meer hebben (of beter is dat hier dan de specifieke waarde voor 'niet van toepassing' vanzelf aan gegeven wordt ['99999'])
Een voorbeeld: Een instelling geeft in eerste instantie aan dat de instelling langdurige zorg levert en vult de indicatoren in die daarop van toepassing zijn. Echter, om een bepaalde reden, wijzigt de instelling het antwoord op de vraag over langdurige zorg in 'nee', nog nadat alle indicatoren ingevuld zijn. Het moet dan niet mogelijk zijn dat de waarden die eerder al waren ingevuld op indicatoren die alleen van toepassing zijn op langdurige zorg gewoon blijven staan.
6. Naast 'geen gegevens beschikbaar' moet ook 'niet van toepassing' mogelijk zijn om aan te geven in de dataset. Zorg dat als een indicator niet van toepassing is op een instelling, dit in de indicatorwaarde wordt aangegeven met een specifieke waarde. Hierbij is het handig dat er verschillende codes zijn voor de verschillende redenen van niet-aanleveren. Denk aan: 999999 indien een indicator niet van toepassing is door de routing (instelling kan niet aanleveren omdat de indicator geen betrekking heeft op het type instelling / soort zorg dat instelling levert / etc.), 88888 Indien de instelling aangegeven heeft dat er geen gegevens beschikbaar zijn.
Codeer deze waarden als missing in het databestand.