



Ontwikkeling van een methodiek om de kwaliteit van aangeleverde data te beoordelen

Deze methodiek is toegespitst op de betrouwbaarheid van het registratieproces en de statistische betrouwbaarheid van de indicatorwaarden

11 augustus 2009

Inhoudsopgave

Voorwoord	4
Hoe deze notitie te lezen?	5
1. Inleiding	6
2. Begrippendefinitie	9
3. Indicatoren	11
4. Uitwerking dimensies beoordeling kwaliteit	13
4.1 Betrouwbaarheid van het registratieproces	13
4.2 Statistische betrouwbaarheid	17
5. Illustratie	21
6. Afsluitende opmerkingen	25
Bijlagen	
A. Expertbijeenkomsten	28
B. Dimensies validiteit en vergelijkbaarheid	31
C. Procesbeschrijving en rapportage beoordeling	33
D. Mogelijk hulpmiddel – definiëring indicator	36

Voorwoord

Zichtbare Zorg verzamelt gegevens die informatie geven over de kwaliteit van zorg geleverd door zorgaanbieders. Deze notitie beschrijft een methodiek, waarmee de kwaliteit van aangeleverde gegevens kan worden beoordeeld. Deze is primair bedoeld voor die organisaties die deze indicatorwaarden gebruiken voor het maken van publieksinformatie of zorginkoopinformatie. Het gaat hierbij dus niet om een uitspraak over de kwaliteit van de zorg.

De beoordeling van indicatorwaarden in juiste perspectief geplaatst

Het werken met kwaliteitsindicatoren staat in de zorg relatief nog in de kinderschoenen en is daarom een zoektocht. Gaandeweg leren we hoe we ermee moeten werken, door steeds nieuwe inzichten te verzamelen, te evalueren en indicatoren te verbeteren. In de tussentijd vraagt de samenleving echter wel om informatie over de kwaliteit van de zorg.

Daarom is het de komende jaren noodzakelijk om de gebruikers van de indicatorwaarden zo goed mogelijk te informeren over de beperkingen die voorlopig aan het gebruik van indicatoren en indicatorwaarden verbonden zijn. Dit doen we met de wetenschap dat deze beperkingen van jaar op jaar aantoonbaar kunnen af nemen. In het project Zichtbare Zorg Ziekenhuizen is daarom ook nadrukkelijk sprake van een groeipad. Dit betekent dat zowel de indicatoren als de indicatorwaarden (en de daarbij behorende beoordeling van de kwaliteit ervan) in de komende jaren worden verbeterd en verfijnd.

De beoordelingsmethodiek die voor u ligt is gebaseerd op de "state of the art" op dit moment en is in samenwerking met verschillende experts ontwikkeld (bijlage A). Deze eerste versie van de methodiek zal worden toegepast op de gegevens aangeleverd in 2009 die gaan over de zorg die in 2008 is geleverd. De evaluatie van deze toepassing zal belangrijke input opleveren voor een tweede versie van de beoordelingsmethodiek. Daarmee vormt deze eerste versie van de methodiek tevens een startpunt voor verdere ontwikkeling. In deze ontwikkeling kan de beoordelingsmethodiek ook gezien worden als een processtap op zichzelf. Door vergelijking van de beoordeling kwaliteit van de aangeleverde data ontstaan prikkels tot verbetering van de betrouwbaarheid van de registratie. De beoordeling vormt daarmee ook een instrument om de beoogde uniformiteit van registratie (en interpretatie) te bevorderen.

Met de in dit voorwoord genoemde nuances is het duidelijk dat het beeld dat nu in 2009 geschetst kan worden over de kwaliteit van de indicatorwaarden een begin is, en niet meer dan een eerste indicatie voor de validiteit, betrouwbaarheid (t.a.v. meting/registratie en statistiek) en vergelijkbaarheid van de gebruikte indicatoren kan zijn.

Hoe deze notitie te lezen?

Hoofdstuk 1, de inleiding, geeft u een beeld over welke indicatoren in 2009 worden aangeleverd binnen Zichtbare Zorg Ziekenhuizen en wat we beogen met het beoordelen van de kwaliteit van de aangeleverde data en indicatorwaarden. Ook wordt daar nogmaals onderstreept dat de methodiek die wordt voorgesteld om de kwaliteit van de aangeleverde data te beoordelen nog in ontwikkeling is. In hoofdstuk 2 zijn de vier kwaliteitsdimensies gedefinieerd: validiteit, betrouwbaarheid van het registratieproces, statistische betrouwbaarheid en vergelijkbaarheid. Daarnaast leest u in dit hoofdstuk waarom deze notitie zich richt op de twee dimensies betrouwbaarheid van het registratieproces en statistische betrouwbaarheid. Validiteit en vergelijkbaarheid worden in bijlage B van deze notitie beschreven. Hoofdstuk 3 presenteert de verdeling van de indicatoren in de huidige indicatorensets over de typen indicatoren: structuur-, proces- en uitkomstindicatoren en beschrijft in welke mate de kwaliteitsdimensies een rol spelen bij de verschillende typen indicatoren.

In hoofdstuk 4 zijn de twee dimensies betrouwbaarheid van het registratieproces en statistische betrouwbaarheid uitgewerkt. Dit hoofdstuk laat zien hoe de beoordelingsmethodiek wordt geconcretiseerd. Hoofdstuk 5 illustreert dit aan de hand van een voorbeeld en beschrijft hoe u de beoordeling terugvindt in de verschillende rapportages en in de database met openbare indicatorwaarden. Hoofdstuk 6 sluit af met een aantal belangrijke opmerkingen, dat de lezer helpt om de beoordelingsmethodiek in de juiste context te plaatsen. Drie opmerkingen zijn het belangrijkste:

- (1) Inzicht in de doelen voor gebruik is heel belangrijk: keuze-informatie stelt andere eisen aan een indicator dan verbeterinformatie. Meerdere doelen zijn gericht op het vergelijken van ziekenhuizen;
- (2) De openbaarheid van de indicatorwaarden: o.a. wordt daar aangegeven hoe de leeswijzer bij de indicatoren er gaat uitzien;
- (3) De verbeterpunten wat betreft de verificatievragen om inzicht te krijgen in de betrouwbaarheid van het registratieproces.

In de laatste alinea van hoofdstuk 2 staat de onderlinge samenhang van de vier dimensies beschreven. Ondanks dat deze notitie zich focust op betrouwbaarheid van het registratieproces en statistische betrouwbaarheid, kan de aanpak op de beoordeling van de validiteit en vergelijkbaarheid niet onbesproken blijven. In bijlage B wordt hier op ingegaan. Juist voor de doelen 'verantwoord openbaar maken' en 'leeswijzer' is een integraal beeld van de vier dimensies belangrijk. In bijlage C leest u hoe voor de meting in 2009 het beoordelingsproces van de kwaliteit van de aangeleverde data en indicatorwaarden is gepland, wanneer welke beslissingen worden genomen en wat er in welke rapportage over de beoordeling van de kwaliteit van de data wordt vermeld.

1. Inleiding

Achtergrond

Ten behoeve van een goede werking van het nieuwe zorgstelsel moeten de belangrijkste veldpartijen in de ziekenhuiszorg de beschikking hebben over kwaliteitsinformatie over ziekenhuizen. De verschillende doeleinden waarvoor de indicatorensets worden gebruikt (meerdere zijn gericht op vergelijking van ziekenhuizen) zijn:

- (1) Keuze-informatie;
- (2) Inkoopinformatie;
- (3) Verwijsinformatie;
- (4) Verantwoordingsinformatie;
- (5) Sturingsinformatie;
- (6) Toezichtinformatie.

Zichtbare Zorg Ziekenhuizen heeft voor tien aandoeningen indicatorensets ontwikkeld, te weten:

- Blaascarcinoom;
- Cataract;
- Diabetes mellitus;
- Heup- en knie vervangingen;
- Incontinentie bij vrouwen;
- Liesbreuk;
- Lumbosacraal Radiculair Syndroom;
- Mammacarcinoom;
- Varices;
- Ziekten van de Amandelen en Tonsillen.

Voor elke aandoening is een aantal indicatoren geformuleerd die gezamenlijk een beeld moeten geven van de kwaliteit van de zorg voor deze aandoening. Er zijn drie typen indicatoren onderscheiden:

- **Structuurindicatoren.** Deze beschrijven of de zorgaanbieder bepaalde voorzieningen heeft getroffen in de organisatie. Bijvoorbeeld: het beschikbaar stellen van een patiëntenfolder of het beschikken over een systeem voor complicatieregistratie. Het betreft veelal het eenmalig beantwoorden van een ja/nee-vraag.
- **Procesindicatoren.** Deze volgen een serie gebeurtenissen tijdens een onderzoek en/of behandeling van een patiënt. Bijvoorbeeld: het percentage operaties dat in dagbehandeling is uitgevoerd of het percentage diabetespatiënten dat in het afgelopen jaar een oogcontrole heeft gehad. Deze indicatoren worden geaggregeerd over de patiënten uitgevraagd (veelal als teller en noemer).
- **Uitkomstindicatoren.** Deze richten zich op de uitkomsten (product/effect) van zorg. Bijvoorbeeld: het percentage patiënten met irradiciteit na eerste (borstbesparende) excisie van een primair mammacarcinoom. Deze indicatoren worden geaggregeerd over de patiënten als teller en noemer uitgevraagd.

Voor de indicatoren die als teller en noemer worden uitgevraagd zijn in de webenquête verificatievragen opgenomen. De verificatievragen geven informatie over het registratieproces. Bij de zorginhoudelijke indicatoren zijn over alle proces- en uitkomstindicatoren verificatievragen gesteld; bij de structuurindicatoren zijn deze vragen niet gesteld.

De 10 indicatorensets worden in 2009 landelijk uitgerold. Op 1 juli 2009 dienen in principe alle ziekenhuizen (ongeveer honderd) middels een webenquête de gegevens voor de indicatoren aan te leveren (inclusief de verificatievragen). Deze gegevens gaan over de zorg die in 2008 is geleverd.

Aanleiding

Het is belangrijk dat de Zichtbare Zorg indicatoren, de aangeleverde data en de indicatorwaarden aan vier kwaliteitseisen voldoen.

- (1) De onderliggende indicatoren dienen valide te zijn;
- (2) De data waarop ze gebaseerd zijn dienen betrouwbaar gemeten en geregistreerd te zijn;
- (3) De indicatorwaarden dienen statistisch betrouwbaar te zijn;
- (4) De indicatorwaarden dienen vergelijkbaar te zijn.

Om dit te monitoren is er behoefte aan een evaluatie van de kwaliteit van de indicatoren, de aangeleverde data en de indicatorwaarden. Daarom ontwikkelt Zichtbare Zorg Ziekenhuizen een systeem om deze kwaliteit te beoordelen. Deze 'Beoordeling' is gebaseerd op de score van de vier eerder genoemde kwaliteitseisen.

Doeleinden voor de beoordeling van kwaliteit

De beoordeling van kwaliteit kan voor verschillende doeleinden worden aangewend. We beogen:

- (1) Verantwoord openbaar maken. Op basis van de score op de vier kwaliteitseisen kan de stuurgroep¹ een onderbouwd besluit nemen over het al dan niet of (deels) openbaar maken van de gegevens. Het uitgangspunt daarbij is dat alle indicatorwaarden openbaar worden samen met een beoordeling voor hun 'kwaliteit'.
- (2) Leeswijzer of disclaimer. De beoordeling geeft de gebruikers (bijvoorbeeld zorginkopers en patiëntenorganisaties) die van de indicatorwaarden informatie moeten maken voor hun doelgroep, inzicht in de mate waarin de indicatoren aan de kwaliteitseisen voldoen. Het is vervolgens de verantwoordelijkheid van deze partijen hoe ze met de kwaliteit in relatie tot de bruikbaarheid van de indicatorwaarden voor hun doeleinden omgaan.
- (3) Optimalisatie. De beoordeling geeft de stuurgroep inzicht in de kwaliteit van de indicatorwaarden. Hierdoor kunnen ze doelmatige keuzes maken bij de acties voor verbetering van de kwaliteit van de indicatoren en de data.
- (4) Monitoring. Aan de hand van de uitkomsten op de beoordeling kan de stuurgroep monitoren of de kwaliteit van de indicatorwaarden verbeteren.

Groeipad

Het werken met indicatoren staat in de zorg relatief nog in de kinderschoenen en is daarom een zoektocht. Gaandeweg leren we hoe we ermee moeten werken, door steeds nieuwe inzichten te verzamelen, te evalueren en indicatoren te verbeteren. In de tussentijd vraagt de samenleving echter wel om informatie over de kwaliteit van de zorg. Het ontwikkelen van deze beoordeling van kwaliteit in een kort tijdsbestek is ambitieus. Het gaat om een nieuwe methodiek, die in samenwerking met experts (bijlage A) uit het veld is ontwikkeld. Met deze brede aanpak hopen we te bewerkstelligen dat het model voor de beoordeling van de kwaliteit van de gegevens breed gedragen wordt en instemming krijgt van de betrokken veldpartijen.

¹ Om de kwaliteit van de sector Ziekenhuizen transparant te maken en kwaliteitsverbetering te stimuleren, hebben acht veldpartijen (NVZ, NFU, Orde, V&VN, ZN, NPCF, Consumentenbond, IGZ) zich verenigd in de Stuurgroep Zichtbare Zorg Ziekenhuizen. De stuurgroep ontwikkelt, implementeert en borgt de indicatoren waarmee de kwaliteit van de zorg in de ziekenhuissector inzichtelijk gemaakt wordt. Programmabureau Zichtbare Zorg verzorgt het secretariaat van de Stuurgroep en ondersteunt.

De methodiek wordt ontwikkeld ten behoeve van de meting Zichtbare Zorg Ziekenhuizen in 2009 over de geleverde zorg in 2008. Ondermeer na een evaluatie van de eerste toepassing op de meetuitkomsten voor 2008 zal de methodiek doorontwikkeld worden. Het beeld dat dit jaar geschetst kan worden van de kwaliteit van de indicatorwaarden is een begin, een eerste indicatie voor de validiteit, betrouwbaarheid (t.a.v. meting/registratie en statistiek) en vergelijkbaarheid.

2. Begrippendefinitie

De beoordeling van de kwaliteit van de indicatoren is gericht op de dimensies validiteit, betrouwbaarheid van het registratieproces, statistische betrouwbaarheid en vergelijkbaarheid. Hieronder worden de vier dimensies gedefinieerd en wordt kort toegelicht hoe deze verwerkt worden in de methodiek voor de beoordeling. De dimensies werken anders uit voor de drie typen indicatoren. Dat wordt beschreven in hoofdstuk 3.

- (1) De *validiteit* van een indicator geeft aan in hoeverre de geoperationaliseerde grootheid overeenkomt met het abstracte concept.
Toelichting: Op basis van een analyse van de aangeleverde data zal over de validiteit van de indicatoren geen uitspraak kunnen worden gedaan. De validiteit van een indicator ligt bij de indicatorwerkgroep. De indicatorwerkgroepen hebben op basis van consensus de indicatoren vastgesteld. Een leidraad daarbij is het AIRE-instrument, een methodologisch instrument dat dient als evaluatie- en toetsingskader voor indicatoren. Hiermee is de indrukvaliditeit van de indicator in principe gewaarborgd. Over de inhouds-, construct- en criteriumvaliditeit kan op basis van de data die door de ziekenhuizen zijn aangeleverd, geen uitspraak worden gedaan². Hiermee zal voor de dimensie validiteit in het huidige rapport geen methodiek voor beoordeling worden ontwikkeld. Wel zullen de opmerkingen van de ziekenhuizen (logboek) die betrekking hebben op validiteit worden geïnventariseerd. In bijlage B is een voorstel voor beoordeling en aanbeveling opgenomen.
- (2) De *betrouwbaarheid van het registratieproces* wordt weerspiegeld in het 'juist, volledig en op tijd' registreren en aanleveren van gegevens. Gegevens zijn juist, wanneer een ziekenhuis *daadwerkelijk* de informatie heeft aangeleverd die in de webenquête is gevraagd: de aangeleverde waarde is gelijk aan de werkelijke onderliggende waarde. De gegevens zijn volledig wanneer een ziekenhuis *alle* informatie die in de webenquête is gevraagd over de *gehele tijdsperiode* heeft aangeleverd. De gegevens zijn tijdig wanneer ze *binnen de gestelde tijdsperiode* door de ziekenhuizen worden aangeleverd.
Toelichting: Op basis van de verificatievragen die zijn opgenomen in de webenquête, zal in dit rapport voor de betreffende indicatoren een methodiek voor de beoordeling van de kwaliteit worden ontwikkeld. Deze wordt besproken in hoofdstuk 4.
- (3) Een indicatorwaarde is *statistisch betrouwbaar* wanneer deze waarde nauwkeurig is gemeten: dit betekent dat bij herhaalde metingen onder dezelfde omstandigheden de indicatorwaarden weinig van elkaar verschillen, er is weinig variatie. Hierbij is de nauwkeurigheid van de individuele waarde van belang als ook de nauwkeurigheid van de vergelijking tussen ziekenhuizen van individuele waarden. Uitgaande van een volledige telling (voorgeschreven in de gidsen) is de statistische betrouwbaarheid in principe niet door het ziekenhuis te beïnvloeden en heeft in die zin geen morele betekenis voor het ziekenhuis. Wanneer een ziekenhuis zijn cijfers baseert op een steekproef, gaat dit niet op.
Toelichting: De beoordeling van deze dimensie zal in hoofdstuk 4 verder worden uitgewerkt.

² In een onderzoek van de NFU zal voor de indicatorensets mammacarcinoom en heup/knievervangende validiteit nader worden onderzocht (uitgevoerd door o.a. prof. Kievit, prof. Klazinga en prof. Steyerberg).

(4) Bij de *vergelijkbaarheid* tussen indicatorwaarden spelen twee aspecten een rol:

(i) Vergelijkbaarheid tussen ziekenhuizen

Verstorende factoren: Bij de vergelijking van indicatorwaarden speelt de invloed van eventuele factoren die verschillen tussen de ziekenhuizen een rol. Veelal gaat het om kenmerken van de patiënten die in een ziekenhuis worden behandeld (leeftijd, comorbiditeit). Deze factoren kunnen van invloed zijn op de waarde van de indicator. Verschillen tussen ziekenhuizen zouden hierdoor (gedeeltelijk) toegeschreven kunnen worden aan verschillen tussen deze factoren. In de vergelijking tussen de ziekenhuizen zou hiermee rekening gehouden kunnen worden door te corrigeren voor deze factoren, een zogenaamde case-mix adjustment. Daarnaast kan men ook denken aan kenmerken van het ziekenhuis zelf³. Het lijkt niet relevant om voor deze ziekenhuiskenmerken te corrigeren, wanneer men de kwaliteit van zorg wil vergelijken. Ze zijn wel relevant om gevonden verschillen tussen ziekenhuizen verder te verklaren.

Toelichting: Momenteel worden er geen case-mix adjusters uitgevraagd. Hierdoor zal er in het huidige rapport geen methodiek worden ontwikkeld om de vergelijkbaarheid tussen ziekenhuizen te beoordelen. Wel is in bijlage B een aantal opmerkingen over en aanbevelingen voor de vergelijkbaarheid toegevoegd.

(ii) Vergelijkbaarheid in de tijd: Binnen een ziekenhuis zijn twee indicatorwaarden vergelijkbaar in de tijd, wanneer de gegevens waarop de indicatoren zijn gebaseerd dezelfde betekenis hebben (bijvoorbeeld: de definitie is niet veranderd, de manier van verzamelen is niet veranderd) en betrekking hebben op een vergelijkbare patiëntengroep (vergelijkbaar voor wat betreft de case-mix adjusters).

Toelichting: Aangezien de indicatorensets dit jaar voor het eerst landelijk zijn uitgevraagd, is een vergelijking van ziekenhuizen niet aan de orde. In het huidige rapport zal daarom geen methodiek worden ontwikkeld om de vergelijkbaarheid in de tijd te beoordelen. Wel is in bijlage B een aantal opmerkingen over en aanbevelingen voor de vergelijkbaarheid toegevoegd.

De volgorde van de dimensies is niet arbitrair: wil een indicator bruikbaar zijn voor het meten van een aspect van de kwaliteit van de geboden zorg, dan moet deze indicator valide zijn. Dit is een eerste vereiste. Wanneer een indicator valide is, dan moet deze betrouwbaar worden geregistreerd, dat wil zeggen volledig en juist. Als een indicator betrouwbaar is geregistreerd, dan moet deze zo nauwkeurig mogelijk gemeten worden, dat wil zeggen statistisch betrouwbaar. Vervolgens kunnen indicatorwaarden tussen en binnen ziekenhuizen worden vergeleken, waar noodzakelijk na toepassing van een correctie met case-mix adjusters.

Een indicatorwaarde kan betrouwbaar geregistreerd en statistisch betrouwbaar gemeten zijn; wanneer de indicator echter niet valide is, dan is de betreffende indicatorwaarde niet informatief.

³ Kenmerken van ziekenhuizen kunnen o.a. zijn: personeel (aanwezigheid, kundigheid), goederen (aanwezige vaste of roerende, zoals werkende medische apparatuur en hulpmiddelen), informatievoorziening (organisatie volgens/kennis van zorgpaden, beschikbaarheid gegevens op de poli) en financiële middelen.

3. Indicatoren

Zoals aangegeven in hoofdstuk 1 zijn er drie typen indicatoren: structuur-, proces- en uitkomstindicatoren. In de indicatorenengidsen staat bij iedere indicator vermeld tot welk type de betreffende indicator wordt gerekend. Voor elke aandoening staat in onderstaande tabel het aantal indicatoren per type samengevat.

Tabel 1. Per aandoening het aantal structuur-, proces en uitkomstindicatoren

Aandoening	Type indicator		
	Structuur	Proces**	Uitkomst
Blaascarcinoom	3	2	
Cataract	8*	3	
Diabetes	12*	5	
Heup/Knie vervanging	18*	4	1
Incontinentie bij de vrouw	6*	2	1
Liesbreukoperaties	2	2	1
Lumbosacraal radiculair syndroom	2	1	1
Mammacarcinoom	14*	4	2
Varices	5*	2	
Ziekten van adenoïd en tonsillen	1	4	2

*alle klantpreferentie indicatoren zijn aangemerkt als structuurindicatoren

**de meeste procesindicatoren zijn uitgevraagd als teller en noemer

De vier eerder genoemde kwaliteitsdimensies (validiteit, betrouwbaarheid registratieproces, statistische betrouwbaarheid en vergelijkbaarheid) werken niet hetzelfde uit voor de drie typen indicatoren (zie tabel 2):

- **Structuurindicator.** Validiteit is bij dit type indicator relevant. Mits goed gedefinieerd èn eenduidig geïnterpreteerd, zullen deze indicatoren betrouwbaar worden geregistreerd (zie ook opmerking geplaatst bij tabel 2). De meeste structuurindicatoren worden uitgevraagd als een eenmalig ja/nee: het bepalen van de statistische betrouwbaarheid op het niveau van de ziekenhuizen is daarom niet aan de orde. De structuurindicatoren zullen goed vergelijkbaar zijn tussen de ziekenhuizen, aangezien verschillen in patiëntenpopulatie geen rol spelen. Mits de definitie van een indicator onveranderd blijft, zullen de structuurindicatoren goed vergelijkbaar zijn in de tijd (binnen ziekenhuizen).
- **Procesindicator.** Ook bij dit type indicator is validiteit relevant. Voor de meeste procesindicatoren zijn ook betrouwbaarheid van het registratieproces en statistische betrouwbaarheid relevante dimensies om te beoordelen. Verschillen tussen patiëntenpopulaties zullen van beperkte invloed zijn wanneer de procesindicatoren tussen de ziekenhuizen worden vergeleken; de beoordeling van de mate van vergelijkbaarheid tussen ziekenhuizen is hiermee relevant voor de procesindicatoren. Bij gelijkblijvende definitie van de indicator, zullen de procesindicatoren goed vergelijkbaar zijn in de tijd (binnen ziekenhuizen)
- **Uitkomstindicator.** Bij dit type indicator zijn alle dimensies belangrijk.

Tabel 2. Relevantie van kwaliteitsdimensies bij de type indicatoren*

	Validiteit	Betrouwbaarheid registratieproces**	Statistische betrouwbaarheid	Vergelijkbaarheid	
				Tussen ziekenhuizen	Binnen ziekenhuizen
Structuurindicator	+	+/-	-	-	-
Procesindicator	+	+/-	+/-	+/-	-
Uitkomstindicator	+	+	+	+	+

* Een 'min' in deze tabel betekent dat voor het type indicator geen rekening gehouden hoeft te worden met deze dimensie. De resultaten zijn in ieder geval betrouwbaar en of vergelijkbaar.

** Belangrijke voorwaarden voor een betrouwbaar registratieproces is een goede definitie van de indicator, een eenduidige interpretatie van de definitie en daarmee samenhangend draagvlak voor de definitie en de interpretatie. De enige manier om goed te onderzoeken of de definitie goed en de interpretatie eenduidig is, is een werkelijk onderzoek op werkvloer noodzakelijk: een vergelijking van hoe de werkelijke situatie er uit ziet met wat wordt ingevuld. Nu geven alleen de verificatievragen (voor proces- en uitkomstindicatoren) en de opmerkingen die ziekenhuizen hebben gemaakt bij de ingevulde indicatorwaarde een aanknopingspunt. Er zijn ideeën voor een dergelijk onderzoek, waarbij de werkelijke situatie vergeleken wordt met de ingevulde waarde.

4. Uitwerking dimensies beoordeling kwaliteit

Op grond van de beschreven dimensies in hoofdstuk 2 maken we een beoordelingsmethodiek die de kwaliteit van de gemeten indicator weergeeft. Deze kwaliteitsmaat wordt voor de dimensies *betrouwbaarheid van het registratieproces* en *statistische betrouwbaarheid* uitgewerkt. In bijlage B staan aanbevelingen voor de kwaliteitsdimensies validiteit en vergelijkbaarheid.

4.1 Betrouwbaarheid van het registratieproces⁴

Deze dimensie zal worden beoordeeld aan de hand van de antwoorden op de verificatievragen. Een voorwaarde voor een betrouwbaar registratieproces is een goede definitie van de indicator die bovendien eenduidig wordt geïnterpreteerd. Structuurindicatoren zullen, mits aan deze voorwaarde is voldaan, betrouwbaar worden geregistreerd. Deze kwaliteitsdimensie speelt dan ook vooral een rol bij proces- en uitkomstindicatoren. De verificatievragen zijn voor een groot aantal proces- en uitkomstindicatoren – te weten de indicatoren die als teller en noemer zijn uitgevraagd – opgenomen in de webenquête. Daarmee kan een inschatting gemaakt worden van de juistheid en volledigheid van de aangeleverde gegevens. Per indicator zal op basis van de informatie uit een deel van de verificatievragen de betrouwbaarheid van het registratieproces worden beoordeeld. De verificatievraag die vraagt naar de duidelijkheid van de definitie is losgekoppeld van de beoordeling van het registratieproces, omdat het een randvoorwaarde is die niets zegt over de inrichting van het registratieproces in een ziekenhuis. Per indicator zal worden aangegeven in hoeverre duidelijk is welke gegevens moeten worden aangeleverd (definiëring indicator). In de spiegelrapportages vindt een terugkoppeling plaats naar de ziekenhuizen. Zij krijgen hiermee inzicht hoe betrouwbaar zijzelf de data voor een indicator hebben geregistreerd en hoe zij zich verhouden ten opzichte van alle ziekenhuizen.

Er dient opgemerkt te worden dat het hierbij gaat om zelfrapportage: de verificatievragen geven de opvatting van de ziekenhuizen zelf over de aangeleverde gegevens weer. Dit brengt een tweetal risico's met zich mee: (1) ziekenhuizen kunnen in de veronderstelling zijn de gegevens voor een indicator juist en volledig te hebben aangeleverd, waarbij onduidelijkheden niet zijn herkend/onderkend; (2) een ziekenhuis kiest sociaal wenselijke antwoordcategorieën. Een betere manier om goed zicht te krijgen om het registratieproces, is een werkelijke audit op werkvloer: een vergelijking tussen de werkelijke situatie en wat er wordt ingevuld. Deze audit wordt ingepland na de meting, in het najaar van 2009.

⁴ Met de term 'registratieproces' wordt het proces binnen een ziekenhuis bedoeld waarin verschillende stappen te onderscheiden zijn: meten, registreren, intern rapporteren (aan bijvoorbeeld medisch specialisten of Raad van Bestuur) en het aanleveren.

Hieronder worden eerst de verificatievragen beschreven. Daarna volgt een voorstel voor de beoordeling, die slechts gebaseerd wordt op een deel van de verificatievragen. De verificatievragen richten zich op zes aspecten:

(1) **Definiëring indicator** (verificatievraag 1)

De verificatievraag controleert in hoeverre de definitie van de te leveren gegevens duidelijk is: 'Is het voor u duidelijk welke gegevens u moet opleveren?'

Een indicator kan alleen maar betrouwbaar gemeten/geregistreerd/aangeleverd worden wanneer de gehanteerde definitie van de te leveren informatie waarop de indicator is gebaseerd, duidelijk is en slechts op één manier – die voor iedereen gelijk is – is te interpreteren. Echter, wanneer ziekenhuizen in deze verificatievraag aangeven dat de definitie voor hen duidelijk is, betekent dit niet dat de definitie eenduidig is⁵. Het percentage ziekenhuizen dat aangeeft de definitie (redelijk) onduidelijk te vinden lijkt beter bruikbaar. Opgemerkt dient te worden dat ziekenhuizen deze vraag mogelijk sociaal wenselijke beantwoorden.

(2) **Gegevensbron** (verificatievraag 2)

'Uit welke bron of bronnen heeft u de gegevens gehaald (bij teller of noemer die is samengesteld uit gegevens van meerdere bronnen zijn meerdere antwoorden mogelijk)?'

In het rapport 'Voorkomen is beter dan genezen' (zie website Zichtbare Zorg: <http://www.zichtbarezorg.nl/page/Programma-Zichtbare-Zorg/Documenten>), een adviesrapport over betrouwbaarheid van meten en registreren, staat dat 'automatisering van de registratie de foutenkans rechtstreeks reduceert'. Dit betekent dat de betrouwbaarheid van de gegevens groter is wanneer deze uit ICT-bronnen zijn gehaald. Hierbij dienen twee kanttekeningen te worden geplaatst:

- (a) Gegevens uit ICT-bronnen zijn alleen betrouwbaar wanneer er een controle op de invoer heeft plaatsgevonden.
- (b) Gegevens uit schriftelijke bronnen zijn niet per definitie minder betrouwbaar.

(3) **Gegevensverzameling** (verificatievraag 3)

'Welke methode heeft u gehanteerd om tot gegevens te komen?'

In de registratiegidsen wordt voorgeschreven om gegevens aan te leveren over een volledige telling. Gegevens zijn volledig en dus in die zin betrouwbaar wanneer ze gebaseerd zijn op een volledige telling. Extrapolatie op basis van een steekproef uit de gegevens geeft een schatting. Een steekproef, mits aselekt getrokken uit gegevens over de gehele periode, geeft een goede schatting van de indicatorwaarde. Wanneer er extrapolatie plaatsvindt op basis van gegevens van een beperkte periode of van een voorgaande periode, dan kan de schatting van de indicatorwaarde de werkelijke waarde mogelijk vertekend weergeven.

⁵ Aan de hand van histogrammen van de geobserveerde indicatorwaarden kan bekeken worden of ziekenhuizen in groepen uiteen vallen (zie figuur 2 in Bijlage D). Is dit het geval, dan lijken de ziekenhuizen de aan te leveren informatie verschillend te interpreteren en is terugkoppeling naar de indicatorwerkgroep noodzakelijk.

- (4) **Gegevensverificatie** (verificatievraag 4a)
 'Is een medisch specialist betrokken bij de door u opgeleverde gegevens?'
 Gevraagd wordt naar de betrokkenheid van een medisch specialist bij de oplevering van de gegevens. Wanneer een medisch specialist de gegevens heeft geautoriseerd of wanneer ze zijn besproken met de medisch specialist, dan zouden deze beter betrouwbaar kunnen zijn dan wanneer dit niet het geval is.
- (5) **Gegevensgebruik** (verificatievraag 4b)
 'Worden de opgeleverde gegevens in uw eigen ziekenhuis gebruikt?'
 In deze vraag kan worden aangegeven of de opgeleverde gegevens in het ziekenhuis zelf worden gebruikt als managementinformatie of voor kwaliteitsverbetering. De gegevens zijn mogelijk beter betrouwbaar wanneer ze inderdaad intern worden gebruikt. De antwoordcategorie 'nee' is wellicht indicatief.
- (6) **Eigen oordeel over geleverde gegevens** (verificatievraag 5)
 'Wat is uw oordeel over de juistheid van de door u aangeleverde gegevens?'
 Gegevens zijn beter betrouwbaar naarmate de leverancier van de gegevens zelf aangeeft dat de juistheid van de aangeleverde gegevens 'goed' is. Ook deze vraag beantwoorden ziekenhuizen mogelijk sociaal wenselijk.

Voorstel voor beoordeling:

De beoordeling is op basis van input van de experts tot stand gekomen. Een beknopte beschrijving van dit proces en de gehanteerde methode staat in bijlage A.

Beoordeling op basis van 3 aspecten

De beoordeling van de betrouwbaarheid van het registratieproces zal gebaseerd worden op de aspecten gegevensverzameling, gegevensverificatie en gegevensgebruik. Om tot een beoordeling van de betrouwbaarheid van het registratieproces te komen wordt per ziekenhuis voor deze drie aspecten vastgesteld of de aangeleverde gegevens voldoen aan de *gewenste antwoordcategorie*. De gewenste antwoordcategorie is de categorie die een goede betrouwbaarheid weerspiegelt. Voor de drie aspecten geldt als gewenste antwoordcategorie:

- Aspect 3: **gegevensverzameling**: 'volledige telling';
- Aspect 4: **gegevensverificatie**: 'autorisatie door medisch specialist';
- Aspect 5: **gegevensgebruik**: 'managementinformatie' en 'kwaliteitsverbetering'.

De afgeleide beoordelingsmaat is gebaseerd op het aantal keren dat een gewenste antwoordcategorie is gekozen bij deze drie aspecten. Niet elk aspect heeft hetzelfde belang in de beoordeling van de betrouwbaarheid van het registratieproces: gebaseerd op de 'Analytical Hierarchy Process' methode van Saaty is bepaald dat de aspecten gegevensverzameling en gegevensverificatie even belangrijk zijn en beiden een keer zo belangrijk dan gegevensgebruik (bijlage A). Dit leidt tot de volgende beoordelingsmaat:

$$\text{beoordelingsmaat} = 4 \times GA_{\text{gegevensverzameling}} + 4 \times GA_{\text{gegevensverificatie}} + 2 \times GA_{\text{gegevensgebruik}}$$

Hierbij is GA gelijk aan 1 wanneer een ziekenhuis de gewenste antwoordcategorie voor het betreffende aspect (verificatievraag) heeft aangegeven en 0 wanneer dit niet zo is. Een ziekenhuis heeft hiermee een minimale score van 0 en een maximale score van 10.

Deze beoordelingsmaat is als volgt geënclassificeerd:

- Wanneer een ziekenhuis een score heeft van 8 of 10, dan krijgt dit ziekenhuis het vlaggetje 'groen': aan de belangrijke voorwaarden voor het betrouwbaar inrichten van het registratieproces wordt (bijna) volledig voldaan.
- Wanneer een ziekenhuis een score heeft van 6, dan krijgt dit ziekenhuis het vlaggetje 'oranje': aan de belangrijke voorwaarden voor het betrouwbaar inrichten van het registratieproces wordt slechts deels voldaan.
- Wanneer een ziekenhuis een score heeft van 0, 2 of 4, dan krijgt dit ziekenhuis het vlaggetje 'rood': er zijn vraagtekens te zetten bij de inrichting van het registratieproces.

De beoordeling van de betrouwbaarheid van het registratieproces voor de indicator -voor alle ziekenhuizen gezamenlijk- wordt gebaseerd op de vlaggetjes per ziekenhuis:

Stap 1:

- Wanneer 25% of meer van de ziekenhuizen het vlaggetje 'rood' heeft, dan krijgt de betrouwbaarheid van het registratieproces voor de betreffende indicator het vlaggetje 'rood': alle ziekenhuizen bezien, zijn er vraagtekens te zetten bij de inrichting van het registratieproces.

Stap II: wanneer minder dan 25% van de ziekenhuizen het vlaggetje 'rood' heeft, dan

- Krijgt de indicator het vlaggetje 'groen' wanneer 75% van de ziekenhuizen het vlaggetje 'groen' heeft: alle ziekenhuizen bezien, wordt aan de belangrijke voorwaarden voor het betrouwbaar inrichten van het registratieproces (bijna) volledig voldaan;
- Krijgt de indicator het vlaggetje 'oranje' in alle andere gevallen: alle ziekenhuizen bezien, wordt aan de belangrijke voorwaarden voor het betrouwbaar inrichten van het registratieproces slechts deels voldaan.

Overige aspecten

Naast bovenstaande beoordeling zal voor elke indicator aangegeven worden wat de verdeling is over de antwoordcategorieën van de drie aspecten die niet meegenomen worden:

- Aspect 1: definiëring indicator. De antwoorden op deze verificatievraag zeggen namelijk niets over het registratieproces in een ziekenhuis.
- Aspect 2: gegevensbron. Het blijkt in dit stadium niet mogelijk om op basis van de verificatievragen een onderscheid te maken in de betrouwbaarheid van de registratiesystemen die worden uitgevraagd. Deze vraag zal niet worden opgenomen in de beoordeling; ter evaluatie zal per indicator een overzicht van de frequenties per gebruikt registratiesysteem worden opgesteld. Het risico op sociaal wenselijke antwoorden bij het eigen oordeel over de geleverde gegevens wordt als te groot gezien om dit aspect in de beoordeling op te nemen. Evaluerend zal er per indicator en per ziekenhuis een overzicht van de frequenties per antwoordcategorie worden opgesteld.
- Aspect 6: eigen oordeel geleverde gegevens. Deze zal uitsluitend op indicatorniveau (dus niet herleidbaar op ziekenhuisniveau) worden geanalyseerd.

4.2 Statistische betrouwbaarheid

Deze kwaliteitsdimensie speelt alleen een rol bij de proces- en uitkomstindicatoren, die veelal als een proportie worden uitgevraagd (ja/nee-vragen, die leiden tot een teller – bijvoorbeeld het aantal keren ‘ja’ – en een noemer – het totale aantal patiënten waarop de vraag betrekking heeft).

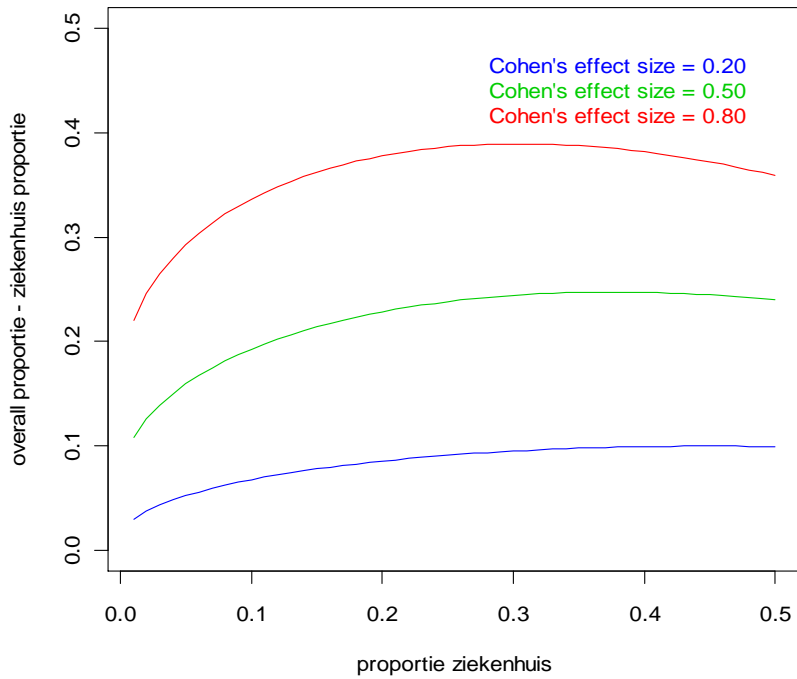
De nauwkeurigheid van een indicatorwaarde (puntschatting) wordt weergegeven door het 95%-betrouwbaarheidsinterval (intervalschatting). Een beoordeling van de statistische betrouwbaarheid zou op basis van dit 95%-betrouwbaarheidsinterval kunnen worden gedaan. Echter, in de inleiding staat dat voor de meeste doeleinden een vergelijking tussen de ziekenhuizen zal worden gemaakt, waarbij men veelal geïnteresseerd zal zijn in relatieve risico's (de proportie in ziekenhuis 1 is twee keer zo groot dan de proportie over alle ziekenhuizen) en niet in absolute verschillen (de proportie in ziekenhuis 1 is 0.10 hoger dan de proportie over alle ziekenhuizen).

Om het totaal aantal benodigde waarnemingen dat nodig is om betrouwbare vergelijkingen te doen te bepalen is het noodzakelijk:

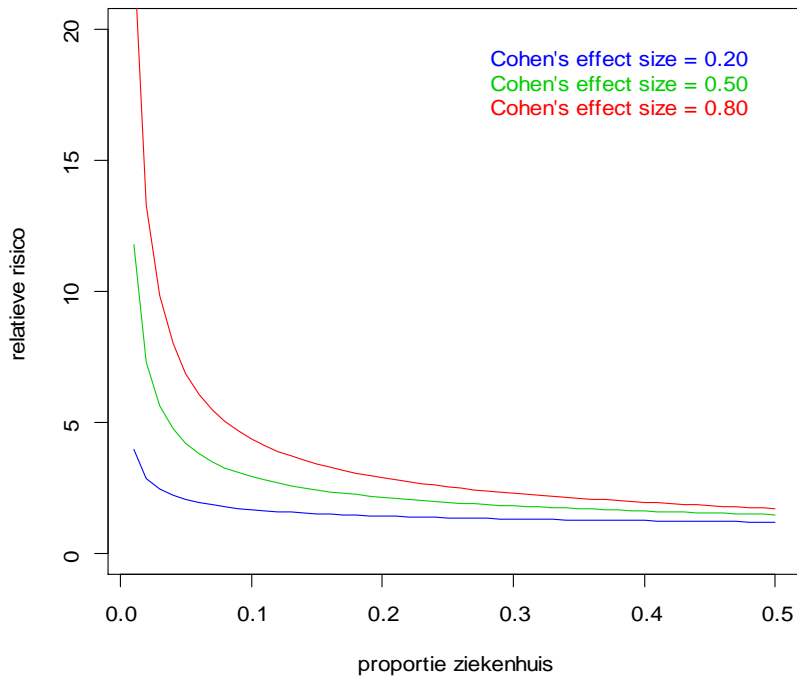
- (1) Aan te geven welk effect als relevant wordt gezien;
- (2) Met welke kans dit effect moet worden gedetecteerd (power of onderscheidend vermogen);
- (3) Met welke onbetrouwbaarheid moet worden vergeleken.

In deze notitie zal de beoordeling van statistische betrouwbaarheid gebaseerd worden op vergelijkingen tussen de ziekenhuizen en is gekozen voor een power van 80% en een onbetrouwbaarheid van 5%. Dit zijn algemeen geaccepteerde waarden.

Om de grootte van het relevant geachte effect vast te stellen, wordt vaak gebruik gemaakt van Cohen's effect size. Kleine effecten komen overeen met een Cohen's effect size van 0.2, middelmatige effecten met 0.5 en grote effecten met 0.8. Hoe kleiner het effect, des te meer waarnemingen er nodig zijn. Cohen's effect size is geen absolute maat (zie figuur 1): hoe kleiner de proportie van het ziekenhuis, des te kleiner het absolute verschil tussen de proportie van het ziekenhuis en de proportie over alle ziekenhuizen. In termen van relatieve risico's is Cohen's effect size constanter (zie figuur 2): vanaf een proportie van 0.05 is voor kleine en middelmatige effecten het relatieve risico constant. Voorgesteld wordt om de beoordeling van de statistische betrouwbaarheid afhankelijk te maken van Cohen's effect size en de grootte van de proportie waargenomen in een ziekenhuis.



Figuur 1. Het absolute verschil tussen de proportie over alle ziekenhuizen en de proportie van een ziekenhuis tegen de proportie van een ziekenhuis voor verschillende waarden van Cohen's effect size.



Figuur 2. Relatieve risico tegen de proportie van een ziekenhuis voor verschillende waarden van Cohen's effect size.

Voorstel beoordeling:

Ook de statistische betrouwbaarheid wordt per ziekenhuis beoordeeld met de vlaggetjes 'groen', 'oranje' en 'rood'. Deze beoordeling wordt gebaseerd op de benodigde steekproefgrootte om een betrouwbare vergelijking te kunnen maken.

Bij proporties groter dan of gelijk aan 0.05 (5%), maar kleiner of gelijk aan 0.95 (95%):

- De statistische betrouwbaarheid 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); het ziekenhuis krijgt dan het vlaggetje 'groen' voor statistische betrouwbaarheid;
- De statistische betrouwbaarheid 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); het ziekenhuis krijgt dan het vlaggetje 'rood' voor statistische betrouwbaarheid;
- Ligt het aantal waarnemingen tussen de 30 en 200, dan krijgt het ziekenhuis het vlaggetje 'oranje' voor statistische betrouwbaarheid, de statistische betrouwbaarheid is 'matig' te noemen.

Bij kleiner dan 0.05 (5%) of groter dan 0.95 (95%):

- De statistische betrouwbaarheid 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); het ziekenhuis krijgt dan het vlaggetje 'groen' voor statistische betrouwbaarheid;
- De statistische betrouwbaarheid 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); het ziekenhuis krijgt dan het vlaggetje 'rood' voor statistische betrouwbaarheid;
- Ligt het aantal waarnemingen tussen de 200 en 800, dan krijgt het ziekenhuis het vlaggetje 'oranje' voor statistische betrouwbaarheid, de statistische betrouwbaarheid is 'matig' te noemen.

De relatieve risico's voor een betrouwbare vergelijking zijn door de keuze van deze Cohen's effect sizes vergelijkbaar over de gehele range van proporties.

De beoordeling van de statistische betrouwbaarheid voor de indicator -voor alle ziekenhuizen gezamenlijk- wordt gebaseerd op de vlaggetjes per ziekenhuis:

Stap 1:

- Wanneer 25% of meer van de ziekenhuizen het vlaggetje 'rood' heeft voor een indicator krijgt deze het vlaggetje 'rood'. De statistische betrouwbaarheid voor de betreffende indicator is slecht: 25% of meer van alle ziekenhuizen heeft een statistisch onbetrouwbare waarde. Een zinnige vergelijking tussen de ziekenhuizen is dan ook niet mogelijk.

Stap II: wanneer minder dan 25% van de ziekenhuizen het vlaggetje 'rood' heeft, dan

- Krijgt de indicator het vlaggetje 'groen' wanneer 75% van de ziekenhuizen het vlaggetje 'groen' heeft. De statistische betrouwbaarheid is 'goed': 75% of meer van alle ziekenhuizen heeft een statistisch betrouwbare waarde. Een zinnige vergelijking tussen de ziekenhuizen is dan ook op statistische gronden mogelijk. Echter, het zegt niets over de dimensie vergelijkbaarheid: er kunnen nog case-mix variabelen zijn die verstoring werken.

- Krijgt de indicator het vlaggetje 'oranje' alle andere gevallen. De statistische betrouwbaarheid is 'matig': een zinnige vergelijking met waarden van andere ziekenhuizen vraagt een genuanceerd oordeel.

5. Illustratie

Om de 'Beoordeling' voor de dimensies *betrouwbaarheid van het registratieproces* en *statistische betrouwbaarheid* te illustreren, wordt in deze paragraaf een fictief voorbeeld gepresenteerd.

Aangeleverde data

We nemen de indicator 'Het percentage patiënten met een niet-spierinvasieve tumor, waarbij na een TUR een eenmalige spoeling met een cytostaticum binnen 24 uur werd gegeven' uit de indicatorenset voor blaascarcinoom. De data-aanlevering voor 14 ziekenhuizen zou er dan als volgt kunnen uitzien (deels fictief):

Tabel 3. Voorbeeld (deels fictief) van aangeleverde gegevens voor de indicator 'Het percentage patiënten met een niet-spierinvasieve tumor, waarbij na een TUR een eenmalige spoeling met een cytostaticum binnen 24 uur werd gegeven'

Ziekenhuis	Aantal waarnemingen		Verificatievragen						Eigen oordeel geleverde gegevens
	Teller *	Noemer *	Definiëring indicator	Gegevens bron	Gegevens verzameling	Gegevens verificatie **	Gegevens gebruik ***		
1	20	200	Nee	DBC	Volledige telling	Nee	Nee	Goed	
2	3	21	Ja	Geen	Protocollair	Ja	Ja	Goed	
3	0	55	Ja	Geen	Volledige telling	Nee	Nee	Goed	
4	116	217	Ja	Geen	Volledige telling	Ja	Ja	Goed	
5	11	114	Ja	DBC	Volledige telling	Ja	Nee	Goed	
6	0	247	Ja	DBC	Volledige telling	Ja	Ja	Goed	
7	39	111	Ja	Geen	Volledige telling	Nee	Nee	Goed	
8	0	68	Ja	Geen	Volledige telling	Ja	Ja	Slecht	
9	28	34	Ja	Geen	Protocollair	Ja	Ja	Goed	
10	3	16	Ja	Geen	Volledige telling	Nee	Nee	Slecht	
11	174	183	Ja	DBC	Volledige telling	Ja	Ja	Goed	
12	1	66	Ja	Geen	Volledige telling	Ja	Ja	Goed	
13	96	116	Ja	DBC	Volledige telling	Ja	Ja	Goed	
14	89	141	Ja	DBC	Volledige telling	Nee	Nee	Goed	
totaal	580	1589							

* De getallen zijn afkomstig uit de dataset van de inlooperperiode, de overige gegevens zijn fictief.

** Ja betekent dat de gegevens zijn geautoriseerd, nee betekent dat de gegevens niet zijn geautoriseerd.

*** Ja betekent dat de gegevens worden gebruikt voor managementinformatie of voor kwaliteitsverbetering, nee betekent dat de gegevens hiervoor niet worden gebruikt.

Definiëring indicator

13 van de 14 ziekenhuizen heeft aangegeven het duidelijk is welke gegevens moeten worden opgeleverd ('ja, de definitie is duidelijk'), dit komt overeen met 93%.

Betrouwbaarheid van het registratieproces

De beoordeling van de dimensie *betrouwbaarheid van het registratieproces* is gebaseerd op de verificatievragen gegevensverzameling, gegevensverificatie en gegevensgebruik. Elk ziekenhuis krijgt op basis van de gekozen antwoordcategorieën punten:

- Aspect 3: **gegevensverzameling**: volledige telling – 4 punten;
- Aspect 4: **gegevensverificatie**: autorisatie door medisch specialist – 4 punten;
- Aspect 5: **gegevensgebruik**: voor managementinformatie of voor kwaliteitsverbetering – 2 punten.

Vervolgens wordt per ziekenhuis op basis van het totale aantal punten over de drie aspecten de beoordelingsmaat geclassificeerd:

- aantal punten 8 of 10: het ziekenhuis krijgt voor deze dimensie het vlaggetje 'groen': aan de belangrijke voorwaarden voor het betrouwbaar inrichten van het registratieproces wordt (bijna) volledig voldaan.
- aantal punten is 6: het ziekenhuis krijgt voor deze dimensie het vlaggetje 'oranje': aan de belangrijke voorwaarden voor het betrouwbaar inrichten van het registratieproces wordt slechts deels voldaan.
- aantal punten is 0, 2 of 4: het ziekenhuis krijgt voor deze dimensie het vlaggetje 'rood': er zijn vraagtekens te zetten bij de inrichting van het registratieproces.

Voor de fictieve dataset staat deze beoordeling in tabel 4 weergegeven (pagina 21).

Tabel 4. Beoordeling van de betrouwbaarheid van het registratieproces.

Ziekenhuis	Aantal waarnemingen		Verificatievragen			Beoordeling	
	Teller*	Noemer*	Gegevens verzameling	Gegevens verificatie **	Gegevens gebruik ***	Punten	Vlaggetje
1	20	200	Volledige telling	Nee	Nee	4	rood
2	3	21	Protocollair	Ja	Ja	6	geel
3	0	55	Volledige telling	Nee	Nee	4	rood
4	116	217	Volledige telling	Ja	Ja	10	groen
5	11	114	Volledige telling	Ja	Nee	8	groen
6	0	247	Volledige telling	Ja	Ja	10	groen
7	39	111	Volledige telling	Nee	Nee	4	rood
8	0	68	Volledige telling	Ja	Ja	10	groen
9	28	34	Protocollair	Ja	Ja	6	geel
10	3	16	Volledige telling	Nee	Nee	4	rood
11	174	183	Volledige telling	Ja	Ja	10	groen
12	1	66	Volledige telling	Ja	Ja	10	groen
13	96	116	Volledige telling	Ja	Ja	10	groen
14	89	141	Volledige telling	Nee	Nee	4	rood
totaal	580	1589					rood

Tenslotte is voor de indicator -voor alle ziekenhuizen gezamenlijk- een vlaggetje voor de betrouwbaarheid van het registratieproces vastgesteld: deze is in dit voorbeeld rood, de betrouwbaarheid van het registratieproces van deze indicator is slecht. 5 van de 14 ziekenhuizen (36%) heeft het vlaggetje 'rood' (toegepaste beslisregel: wanneer 25% of meer van de ziekenhuizen het vlaggetje 'rood' heeft, krijgt de indicator het vlaggetje 'rood', zie paragraaf 4.1).

Statistische betrouwbaarheid

De beoordeling van statistische betrouwbaarheid wordt per ziekenhuis bepaald. De toe te passen beslisregels zijn afhankelijk van de geschatte indicatorwaarde over alle ziekenhuizen (zie hoofdstuk 4). In dit voorbeeld is de geschatte indicatorwaarde 36.5% (580/1589; de proportie is dus groter dan 0.05). De beoordeling voor de statistische betrouwbaarheid is als volgt per ziekenhuis:

- Wanneer het totale aantal waarneming in een ziekenhuis meer is dan 200, dan is de statistische betrouwbaarheid van de indicator 'goed' en krijgt het ziekenhuis het vlaggetje 'groen' voor deze dimensie;
- Wanneer het totale aantal waarneming in een ziekenhuis meer is dan 30, maar minder dan 200, dan is de statistische betrouwbaarheid van de indicator 'matig' en krijgt het ziekenhuis het vlaggetje 'oranje' voor deze dimensie;
- Wanneer het totale aantal waarneming in een ziekenhuis 30 is of minder, dan is de statistische betrouwbaarheid van de indicator 'slecht' en krijgt het ziekenhuis het vlaggetje 'rood' voor deze dimensie;

Tenslotte is voor de indicator -voor alle ziekenhuizen gezamenlijk- een vlaggetje voor statistische betrouwbaarheid bepaald: de indicator krijgt het vlaggetje 'oranje', de statistische betrouwbaarheid is matig. Twee van de 14 ziekenhuizen hebben het vlaggetje 'rood' (14.3%); 3 van de 14 ziekenhuizen hebben het vlaggetje 'groen' (21.4%); gebaseerd op deze percentages krijgt de indicator het vlaggetje 'oranje' voor statistische betrouwbaarheid (zie paragraaf 4.2). Voor de fictieve dataset staat de beoordeling voor statistische betrouwbaarheid in tabel 5 weergegeven.

Tabel 5. *Beoordeling van de statistische betrouwbaarheid.*

Ziekenhuis	Teller	Noemer	Vlaggetje
1	20	200	groen
2	3	21	rood
3	0	55	oranje
4	116	217	groen
5	11	114	oranje
6	0	247	groen
7	39	111	oranje
8	0	68	oranje
9	28	34	oranje
10	3	16	rood
11	174	183	oranje
12	1	66	oranje
13	96	116	oranje
14	89	141	oranje
totaal	580	1589	oranje

6. Afsluitende opmerkingen

Tijdens de expertbijeenkomsten is een aantal opmerkingen naar voren gekomen waarvoor in de toekomst doorontwikkeling en/of verdere discussie noodzakelijk is. Deze opmerkingen worden hieronder puntsgewijs besproken.

Doeleinden indicator

Indicatorsets worden voor verschillende doeleinden gebruikt:

- (1) Keuze-informatie;
- (2) Inkoopinformatie;
- (3) Verwijsinformatie;
- (4) Verantwoordingsinformatie;
- (5) Sturingsinformatie;
- (6) Toezichtinformatie.

De eisen die worden gesteld aan de validiteit, betrouwbaarheid (registratieproces en statistisch) en vergelijkbaarheid, zijn niet noodzakelijk hetzelfde voor de verschillende doeleinden. Voor bijvoorbeeld sturings- en toezichtinformatie is de eis van vergelijkbaarheid minder van belang. Zorgverzekeraars, patiëntenorganisaties en verwijzers zullen de indicatorwaarden vooral gebruiken om zorgaanbieders onderling te vergelijken en daarin keuzes te maken: zorgverzekeraars om hun zorginkoop af te stemmen, patiëntenorganisaties om hun achterban gericht te adviseren over de zorgaanbieders.

Gebruikers zullen moeten specificeren welke indicatoren zij willen gebruiken en waarvoor. De noodzakelijke kwaliteit van een indicator en indicatorwaarden hangt hiermee nauw samen: de doelstelling van een indicator (wie gaat de indicator met welk doel gebruiken) bepaalt de noodzakelijke kwaliteit. Dit zal in de toekomst verder worden uitgewerkt.

Gehanteerde dimensies

De relevantie van de vier dimensies waarop de beoordeling van de kwaliteit van aangeleverde data is gericht, staat vast; er is hierover geen discussie. Ook was er consensus over het onderscheid tussen enerzijds de dimensies betrouwbaarheid van het registratieproces en statistische betrouwbaarheid, en anderzijds de dimensies validiteit en vergelijkbaarheid. Voor de eerstgenoemde dimensies kan de beoordeling aan de hand van de aangeleverde data plaatsvinden, voor de laatstgenoemde dimensies ligt de beoordeling meer op het terrein van de indicatorwerkgroep. Deze notitie spitst zich toe op de beoordeling van de aangeleverde data en behandelt daarom voornamelijk de beoordeling van de betrouwbaarheid. De beoordeling van de validiteit en vergelijkbaarheid worden in bijlage B uitgewerkt.

Als mogelijk ontbrekende dimensie wordt discriminerend vermogen genoemd. Het discriminerende vermogen wordt in de ontwikkelfase van een indicator meegenomen als belangrijke voorwaarde voor een indicator en is daarom buiten beschouwing gelaten in de huidige notitie. In de evaluatie van de indicatoren op basis van de aangeleverde data zal er wel aandacht worden geschonken aan het discriminerende vermogen. Opgemerkt dient te worden dat – ondanks een laag discriminerend vermogen van een indicator – een indicator wel van belang kan zijn, omdat het een relevant kwaliteitsaspect meet en/of dat alle ziekenhuizen laag scoren.

Dimensie betrouwbaarheid van het registratieproces

De discussie rond deze dimensie heeft zich op de verificatievragen gericht. De verificatievragen zijn een zelfrapportage van de ziekenhuizen en daarmee subjectief. In de toekomst zou een steekproef uit de ziekenhuizen getrokken kunnen worden die wordt gecontroleerd, of er zou een double-data entry kunnen worden toegepast. Ook wordt er gedacht aan een onderzoek naar de stand van registratie in de ziekenhuizen, waarbij de werkelijke situatie vergeleken wordt met de ingevulde waarde. Hierbij zal de werkelijke situatie op de werkvloer vergeleken worden met wat er wordt ingevuld, op basis van een steekproef uit de ziekenhuizen.

Algemeen wordt opgemerkt dat een aantal verificatievragen volgend jaar scherper geformuleerd kan worden. Specifiek voor de verificatievraag over de gegevensbronnen wordt door de experts aangegeven dat er in de toekomst naartoe moet worden gewerkt om *per indicator* een hiërarchische indeling te maken van gegevensbronnen. De experts geven verder aan dat (een deel van) de verificatievragen ook gesteld zouden moeten worden bij de structuurindicatoren.

Openbaarheid indicatorwaarden

Groeipad

Het werken met indicatoren staat in de zorg relatief nog in de kinderschoenen en is daarom een zoektocht. Gaandeweg leren we hoe we ermee moeten werken, door steeds nieuwe inzichten te verzamelen, te evalueren en indicatoren te verbeteren. In het project Zichtbare Zorg Ziekenhuizen is daarom nadrukkelijk sprake van een groeipad, wat betekent dat zowel de indicatoren als de indicatorwaarden (en de daarbij behorende beoordeling van de kwaliteit ervan) in de komende jaren verbeterd en verfijnd worden. De uiteindelijke doelstelling is te komen tot betrouwbare en valide data die de kwaliteit van de zorg in een ziekenhuis duiden en waarmee vergelijkingen te maken zijn tussen en binnen ziekenhuizen.

Openbaarheid

Het uitgangspunt⁶ is dat alle indicatorwaarden openbaar worden gemaakt. De experts geven aan dat in dit beginstadium er terughoudendheid moet zijn in het naar buiten brengen van indicatorwaarden van met name uitkomstindicatoren, aangezien er nog geen case-mix adjustment wordt toegepast. Het achterhouden van informatie komt de transparantie echter niet ten goede. Daarom wordt het belang van de ontwikkeling van deze beoordelingsmethodiek onderschreven. De data moeten met enerzijds een beoordeling van de betrouwbaarheid van het registratieproces en de statistische betrouwbaarheid en met anderzijds een goede leeswijzer⁷ – waarin de beperkingen worden beschreven – openbaar worden gemaakt. Als suggestie geven de experts mee om op basis van de uitkomsten in de kwantitatieve analyse en evaluatie te komen tot twee openbare databases: een met indicatorwaarden waarvan de kwaliteit voldoende is en een met indicatorwaarden waarvan de kwaliteit nog onvoldoende is. Een scheiding vooraf in wat aangemerkt kan worden als een indicator ‘van voldoende kwaliteit’ of niet is wellicht moeilijk te maken en zou op basis van de uitkomsten van de huidige data kunnen worden vastgesteld. Een mogelijk risico hierbij kan zijn dat de scheiding dan ‘data-driven’ wordt.

⁶ Zoals vastgelegd in de samenwerkingsafspraken van de stuurgroep Zichtbare Zorg Ziekenhuizen, op teller- en noemerniveau.

⁷ De leeswijzer kan gebaseerd worden op de leeswijzer uit de oorspronkelijke etalagegidsen (2006/2007, opgesteld door CBO en Orde).

Leeswijzer

Bij de presentatie van de indicatorwaarden moet, samen met de beoordeling van de in deze notitie beschreven kwaliteitsdimensies, een leeswijzer worden opgenomen. Hierin moet in ieder geval worden beschreven:

- (1) Het 'groeipad' van de indicatorsets en de beoordelingsmethodiek.
- (2) Algemene beperkingen van het gebruik van indicatoren en de gepresenteerde resultaten/openbare data.
- (3) De opmerkingen over de validiteit en vergelijkbaarheid (bijlage B).

Van data naar informatie

Zichtbare Zorg verzamelt gegevens over de kwaliteit van zorg. Deze gegevens zullen moeten worden vertaald naar informatie over de kwaliteit van zorg. Zichtbare Zorg dient hier sturing aan te geven door onder andere het verstrekken van een beoordeling van de kwaliteit van de data en een leeswijzer.

Bijlage A: Expertbijeenkomsten

Insteek bij het ontwikkeltraject van de beoordelingsmethodiek

Het ontwikkelen van deze beoordeling van kwaliteit in een kort tijdsbestek is ambitieus. Het gaat om een nieuwe methodiek, die in samenwerking met experts uit het veld wordt ontwikkeld. Met deze brede aanpak hopen we te bewerkstelligen dat het uiteindelijk te kiezen model voor de beoordeling van de kwaliteit van de gegevens breed gedragen wordt en instemming krijgt van de betrokken veldpartijen.

Tabel A.1: *overzicht van de betrokken experts*

Naam		Voornaam	titel en voorletters	Organisatie	Afdeling/Functie	Aanwezig bij 1 ^e bijeenkomst	Aanwezig bij 2 ^e bijeenkomst
Ahaus		Kees	prof. dr. ir. C.T.B.	TNO Management Consultants en Universiteit Groningen	Hoogleraar kwaliteitsmanagement	Ja	Alleen huiswerk
Bal		Ronald	prof. dr. R	iBMG	Sociaal Medische Wetenschappen	Nee, wel schriftelijke input	Alleen huiswerk
Barneveld	van	Teus	ir. T.A.	Orde		Nee, wel mondelinge input	Ja
Boerrigter		Edmée	mw. drs. E.M.	Medisch Centrum Alkmaar	Project Manager Kwaliteit	Ja	Ja
Duchateau		Quirijn	drs. Q.	Spaarne ziekenhuis	kwaliteitscontroller	nvt	Ja
Hermens		Rosella	mw. dr. R.P.M.G.	IQ healthcare	Senior onderzoeker	Ja	Ja
Király		Thomas	dhr. T.G.R.	Lucas Andreas Ziekenhuis en NVMA (vereniging voor Zorgadministratie en Informatie)		nvt	Ja
Klazinga		Niek	Prof.dr. N.S.	AMC	Sociale Geneeskunde	Ja	Ja
Kroonen		Adri	drs. A., arts	Canisius-Wilhelmina Ziekenhuis	adviseur kwaliteitszorg; afdeling kwaliteit, veiligheid en verantwoording	nvt	Ja
Lei	van der	Johan	Prof. dr. J.	Erasmus MC	Medische Informatica	Ja	Ja
Ouden	den	Lya	mw. dr. A.L.	IGZ		Ja	Nee
Steyerberg		Ewout	prof. dr. E.W.	Erasmus MC	Maatschappelijke Gezondheidszorg	Ja	Ja
Wouters		Michel	drs. M.W.J.M.	AVL/NKI en DSCA (Dutch Surgical Colorectal Audit)		Ja	Ja

Contouren van de beoordelingsmethodiek

In de eerste expertbijeenkomst hebben we de experts (4 bijeenkomsten/gesprekken) gesproken over de uitwerking van de beoordeling aan de hand van een eerste opzet. De opmerkingen van de experts zijn verwerkt in het stuk dat in de junivergadering is voorgelegd aan de stuurgroep Zichtbare Zorg Ziekenhuizen en waarvan zij de contouren van de beoordeling heeft vastgesteld. In de vorige vergadering zijn de contouren zoals beschreven in deze notitie al vastgesteld door de stuurgroep. De grote lijn in deze huidige notitie is dan ook slechts op enkele punten aangescherpt op basis van commentaar uit de expertgroep. Om de leesbaarheid te verbeteren, is aan het begin van het rapport een tekst met aanwijzingen hoe deze notitie te lezen toegevoegd. Tevens is een processchema toegevoegd in bijlage C met daarin de planning in 2009.

Operationalisatie van de dimensies

De operationalisatie van de dimensies is tot stand gekomen in discussie en samenspraak met de experts (2 bijeenkomsten). Ter voorbereiding op de bijeenkomst ontvingen de experts een 'huiswerkopdracht' met daarin:

1. Een enquête over de gewenste antwoordcategorieën voor de verificatievragen. Per mogelijk antwoord op een verificatievraag werd gevraagd aan te geven of de antwoordcategorie gewenst of ongewenst of neutraal wordt gevonden.
2. Een invulopdracht om de weging van de verschillende aspecten van betrouw registreren vast te stellen. Hierbij werd gevraagd om telkens een deelaspect van betrouwbaar registreren te vergelijken met een ander deelaspect: paarsgewijze vergelijking volgens de AHP-methodiek van Saaty. Hierbij moet per vergelijking worden aangegeven welk deelaspect belangrijker wordt gevonden en de mate van belang conform een aangegeven schaling.

Zowel in de uitkomsten van de enquête (tabel A.2) als in de uitkomsten van de paarsgewijze vergelijking (tabel A.3) als in de discussie was er een grote mate van consensus. De experts staan achter de operationalisatie zoals nu gepresenteerd wordt aan de stuurgroep.

Tabel A.3. *Gewichten drie deelaspecten bepaald met behulp van AHP-methodiek (Saaty)*

Aspect	Gewicht
Gegevensverzameling	0.44
Gegevensverificatie	0.40
Gegevensgebruik	0.16

Tabel A.2. Uitkomsten enquête 'gewenste antwoordcategorieën' samengevat voor alle experts

	gewenste antwoord categorie	neutraal	ongewenste antwoord categorie	niet ingevuld
1. Is het voor u duidelijk welke gegevens u moet opleveren?				
a. Ja, de definitie is duidelijk	9			1
b. Redelijk duidelijk, maar er is wel enige ruimte voor interpretatie van de definitie	3	5	1	1
c. Redelijk onduidelijk, er is veel ruimte voor interpretatie van de definitie	2	1	6	1
d. Nee, de definitie is totaal niet duidelijk	2		7	1
2. Uit welke bron of bronnen heeft u de gegevens gehaald?				
a. DBC-registratie	5	2	2	1
b. Verrichtingenregistratie	6	3		1
c. EPD	6	2	1	1
d. specifiek voor de aandoening (* zie onderstaande tabel)	4			6
e. (Poli)klinische status (handmatig)	6		3	1
f. Anders, toelichting	3	4	2	1
<hr/>				
*specifieke bron per aandoening				
Adenoid van de tonsillar Pijn-metingsysteem	7			3
Cataract	6	1		3
Diabetes	5		1	4
Heup/knie	6		1	3
Liesbreuk	7			3
Lumboradicaal Synd	5	2		3
Mammacarcinoom	5	2		3
	5	2		3
	6	1		3
	6		1	3
	6		1	3
	4	2	1	3
	4	3		3
<hr/>				
3. Welke methode heeft u gehanteerd om tot de gegevens te komen?				
a. Volledige telling	9	1		
b. Extrapolatie op basis van beperkte periode	3,5	2,5	4	
c. Extrapolatie op basis van steekproef over de gehele periode	4	5	1	
d. Extrapolatie op basis van gegevens vorige periode	3	1	6	
e. Protocollair, standaard	3	4	3	
f. Anders, toelichting	2	3	3	2
4a. Is een medisch specialist betrokken bij de door u opgeleverde gegevens?				
a. Ja, de gegevens zijn geautoriseerd door een medisch specialist	10			
b. Ja, de gegevens zijn besproken met een medisch specialist	5	2	3	
c. Ja, de resultaten zijn besproken met een medisch specialist	1	7	1	1
d. Nee	3	2	5	
4b. Worden de opgeleverde gegevens in uw eigen ziekenhuis gebruikt?				
a. Ja, voor managementinformatie	7	1	1	1
b. Ja, voor kwaliteitsverbetering	7	1	1	1
c. ja, voor externe verantwoording	2	5	1	2
d. Nee	1	1	7	1
5. Wat is uw oordeel over de juistheid van de door u aangeleverde gegevens?				
a. Goed	6	2	2	
b. Redelijk		7	3	
c. Matig		3	7	
d. Slecht		1	9	

Bijlage B. Dimensies validiteit en vergelijkbaarheid

Deze notitie focust zich op betrouwbaarheid van het registratieproces en statistische betrouwbaarheid aan de hand van de aangeleverde data. De beoordeling van validiteit en vergelijkbaarheid is echter ook van belang, maar ligt meer op het terrein van de indicatorwerkgroep. Juist voor de doeleinden verantwoord openbaar maken en een leeswijzer is een integraal beeld van de vier dimensies belangrijk.

1. Validiteit

Deze dimensie speelt een rol bij alle typen indicatoren: structuur-, proces- en uitkomstindicatoren.

Voor de verschillende indicatorensets geldt over de validiteit het volgende:

- (1) In de indicatorengids staat per indicator aangegeven wat de validiteit van de betreffende indicator is. Gebaseerd op wetenschappelijke publicaties wordt een onderbouwing gegeven voor de validiteit van de indicator.
- (2) De opmerkingen van ziekenhuizen in de webenquête en de vragen die aan de helpdesk zijn gesteld kunnen ook aanwijzingen geven over de validiteit van een indicator.

Voorstel voor beoordeling:

Gebaseerd op de toelichtingen in de indicatorgids over validiteit, aangevuld met een inventarisatie van opmerkingen in de webenquête en vragen die aan de helpdesk zijn gesteld, zal aan de indicatorwerkgroep worden voorgelegd in hoeverre een indicator als valide aan te merken is voor het te meten concept. Indien mogelijk zal hieraan op basis van de expertopinie van de indicatorwerkgroep een vlaggetje 'groen', 'oranje' of 'rood' gegeven worden. Het is daarmee een kwalitatieve beoordeling.

Aanbeveling:

De leden van de indicatorwerkgroep moeten regelmatig bijhouden in hoeverre nieuwe wetenschappelijke inzichten zijn ontstaan, die van invloed zijn op de betreffende indicator. Hernieuwde inzichten kunnen leiden tot aanpassing van de indicator.

2. Vergelijkbaarheid

Deze dimensie speelt alleen een rol bij de uitkomstindicatoren, en mogelijk bij enkele procesindicatoren. Algemeen geldt dus dat voor de structuur- en een deel van de procesindicatoren geen rekening gehouden hoeft te worden met deze dimensie en de resultaten in ieder geval vergelijkbaar zijn. Bij vergelijkbaarheid kan een onderscheid gemaakt worden tussen:

- (1) Vergelijkbaarheid tussen ziekenhuizen: versturende factoren;
- (2) Vergelijkbaarheid in de tijd (binnen ziekenhuizen);

Ad 1

Bij de vergelijking van indicatorwaarden kan de invloed van eventuele cliëntgebonden factoren, die verschillen tussen de ziekenhuizen, een rol spelen. Dit geldt voor de uitkomstindicatoren en voor de procesindicatoren in meer beperkte mate (zie hoofdstuk 3).

Voor de uitkomst- en procesindicatoren zou in de vergelijking tussen de ziekenhuizen rekening gehouden moeten worden met relevante patiëntkenmerken door een zogenaamde case-mix adjustment toe te passen. Hiertoe moet per indicator bepaald worden wat mogelijke case-mix adjusters zijn. Zijn deze case-mix adjusters eenmaal vastgesteld, dan is een goede vergelijking tussen de ziekenhuizen mogelijk. Men moet zich hierbij realiseren dat het moeilijk is om alle mogelijke case-mix adjusters te identificeren, waardoor verschillen gevonden tussen ziekenhuizen na correctie met de geïdentificeerde case-mix adjusters nog steeds veroorzaakt kunnen worden door niet-geïdentificeerde case-mix adjusters.

In de indicatorgidsen staat vermeld in hoeverre de indicatorwerkgroep verwacht dat case-mix adjusters van invloed zijn op de verschillende indicatoren. Nieuwe inzichten in de wetenschappelijke literatuur met betrekking tot mogelijke case-mix adjusters moeten door de leden van de indicatorwerkgroep worden bijgehouden.

Ad 2

Twee indicatorwaarden zijn volledig vergelijkbaar met elkaar, als de gegevens waarop deze indicatorwaarden zijn gebaseerd dezelfde betekenis hebben. Wanneer de indicator is aangepast aan voortschrijdend inzicht in de wetenschappelijke literatuur, dan is vergelijking in de tijd niet meer mogelijk: de betekenis van de indicator is veranderd en geeft hierdoor niet dezelfde informatie in de verschillende jaren. Met de gegevens die de ziekenhuizen dit jaar opleveren is nog geen vergelijking in de tijd mogelijk, aangezien er nog geen informatie van voorgaande jaren beschikbaar is.

Voorstel voor beoordeling:

Gebaseerd op de toelichtingen in de indicatorgids over vergelijkbaarheid zal aan de indicatorwerkgroep worden voorgelegd in hoeverre een indicator als vergelijkbaar aan te merken is. Indien mogelijk zal hier op basis van de expertopinie van de indicatorwerkgroep een vlaggetje 'groen', 'oranje' of 'rood' aangegeven worden. Het is daarmee een kwalitatieve beoordeling.

Aanbevelingen:

- (1) In andere onderzoeken worden leeftijd, geslacht en ernst van de aandoening vaak gebruikt als case-mix adjusters om vergelijking tussen zorgaanbieders mogelijk te maken. De indicatorwerkgroep zou voor vervolgmetingen kunnen overwegen de webenquête uit te breiden met vragen over deze case-mix adjusters. Hierbij moet wel opgemerkt worden dat het voor dit doel wenselijk is dat ziekenhuizen data bij voorkeur op het niveau van patiënten (ook voor de indicatoren) moeten aanleveren. Correctie met case-mix adjusters op geaggregeerd niveau is minder wenselijk, maar verdient de voorkeur boven analyses waarbij case-mix adjusters geheel buiten beschouwing worden gelaten.
- (2) Nieuwe wetenschappelijk inzichten leiden mogelijk tot nieuwe case-mix adjusters. Regelmatige screening van de wetenschappelijke literatuur op dit onderwerp is daarom noodzakelijk.

Bijlage C. Procesbeschrijving en rapportage beoordeling

Juist voor de doeleinden verantwoord openbaar maken en een leeswijzer is een integraal beeld van de beoordeling op vier dimensies validiteit, betrouwbaarheid van het registratieproces, statistische betrouwbaarheid en validiteit belangrijk. Deze bijlage beschrijft in de eerste tabel hoe voor de meting in 2009 het proces van beoordelen van de kwaliteit van de aangeleverde data en indicatorenwaarden is gepland en waar welke beslissingen worden genomen. Vervolgens somt de tweede tabel op wat in welke rapportage/database gerapporteerd wordt over de beoordeling op de vier dimensies.

Tabel C.1: *Proces van de beoordeling van de kwaliteit van de aangeleverde data en indicatorenwaarden*

Stap	Toelichting	Planning
1. Opstellen beoordelingsmethodiek voor de kwaliteit van de aangeleverde data en indicator(waard)en	Zichtbare Zorg en onderzoeks- en adviesbureau Significant stellen in samenwerking met een klankbordgroep van experts (bijlage A) een beoordelingsmethodiek op in 2 fases: 1) Contouren van de beoordelingsmethodiek. 2) Operationalisatie van de dimensies betrouwbaarheid van het registratieproces en statistische betrouwbaarheid: vaststellen hoe maatstaven (kleurcodering groen/oranje/rood) worden toegekend.	Mei / juni 2009
2. Stuurgroep besluit over de uitgewerkte beoordelingsmethodiek	In de stuurgroepvergadering van 4 juni zijn de contouren van de beoordelingsmethodiek vastgesteld (versie 27 mei) en is besloten om de verdere uitwerking vast te stellen in een schriftelijke ronde begin juli.	Juli 2009
3. Data-analyse	Onderzoeks- en adviesbureau Significant voert de data-analyse uit: - Kwantitatieve analyse van de indicatorwaarden. - Beoordeling van de betrouwbaarheid van het registratieproces en statistische betrouwbaarheid volgens de in deze notitie beschreven methodiek. - Voorstel voor leeswijzer en aanwijzingen voor interpretatie.	Juli / augustus 2009
4. Indicatorwerkgroepen	Indicatorwerkgroep: - Presentatie van de beoordeling betrouwbaarheid van het registratieproces en statistische betrouwbaarheid. - Beoordeling indrukvaliditeit en vergelijkbaarheid: als mogelijk worden op basis van een kwalitatieve weging door de werkgroep kleurcodering groen/oranje/rood toegekend. - Leeswijzer wordt besproken en waar nodig aangepast.	Augustus / september 2009

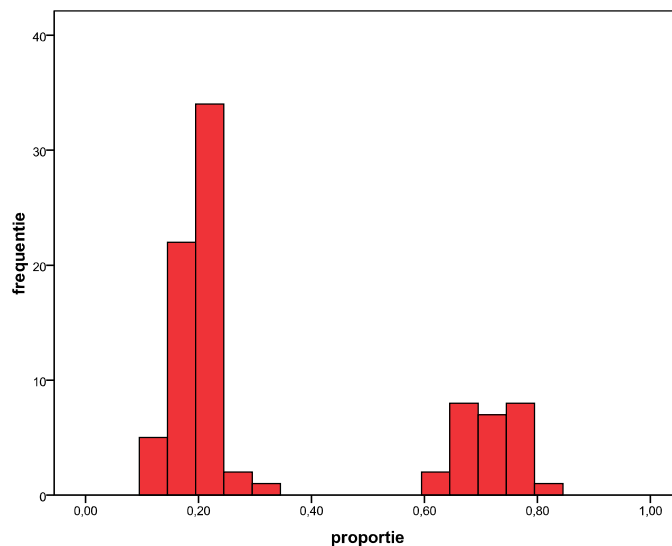
	- Advies over openbaarheid aan de stuurgroep. NB: Zover het mogelijk is binnen de tijdsplanning worden de wetenschappelijke verenigingen betrokken bij/uitgenodigd voor de bijeenkomsten met de indicatorwerkgroepen.	
5. Stuurgroep stelt algemene rapportage vast en besluit over openbaarheid	- Stuurgroep neemt kennis van de algemene rapportage met daarin de kwantitatieve analyse en evaluatie van de indicatoren en de beoordeling op de 4 dimensies (inclusief leeswijzer) en stelt deze rapportage vast. - Stuurgroep besluit over openbaarmaking van de gegevens. Het stuurgroepbesluit over openbaarheid zal gecommuniceerd worden als een zwaarwegend advies aan ziekenhuizen die de definitieve beslissing hebben.	17 september 2009
6. Publicatie van de algemene rapportage	Publicatie van de algemene rapportage op de website van Zichtbare Zorg met daarin de kwantitatieve analyse en evaluatie van de indicatoren en de beoordeling op de 4 dimensies (inclusief leeswijzer). Zie tabel B.2 voor opsomming wat daarin opgenomen is voor de beoordeling op de 4 dimensies.	
7. Spiegelrapportage aan ziekenhuizen	Spiegelrapportage geeft een ziekenhuis inzicht in de eigen data t.o.v. het landelijke beeld (gemiddelde/hogste scores). Zie tabel B.2 voor opsomming wat daarin opgenomen is voor de beoordeling op de 4 dimensies.	
8. Ziekenhuis besluit over de openbaarmaking van de eigen gegevens..	Voor de meting in 2009 hebben de ziekenhuizen de laatste beslissende stem in het wel/niet openbaar beschikbaar stellen van hun gegevens..	16 oktober 2009
9. Publicatie openbare database	Gegevens zijn voor iedereen opvraagbaar middels een verzoek aan Zichtbare Zorg. Zie tabel B.2 voor opsomming wat daarin opgenomen is voor de beoordeling op de 4 dimensies.	Zo spoedig mogelijk na 16 oktober

Tabel C.2: *Wat wordt er in welke rapportage / database gerapporteerd over de beoordeling?*

Rapportage	Wat wordt er gerapporteerd?
<p>Algemene rapportage: Kwantitatieve analyse en evaluatie</p>	<ul style="list-style-type: none"> • De verdeling van de antwoordcategorieën van de verificatievraag over de helderheid van de definitie van de indicator over de ziekenhuizen. • Het vlaggetje van de betrouwbaarheid van het registratieproces, per indicator. • De verdeling van de vlaggetjes voor de betrouwbaarheid van het registratieproces over de ziekenhuizen. • Het vlaggetje van de statistische betrouwbaarheid, per indicator. • De verdeling van de vlaggetjes voor de statistische betrouwbaarheid over de ziekenhuizen. • Uitspraak over de indrukvaliditeit en vergelijkbaarheid door de indicatorwerkgroep (bijlage B). • Per indicator een leeswijzer met een algemene toelichting op het gebruik van indicatoren en voor die indicator specifieke aanwijzingen welke conclusies uit de data kunnen worden getrokken.
<p>Spiegelrapportage aan een ziekenhuis</p>	<ul style="list-style-type: none"> • De verdeling van de antwoordcategorieën van de verificatievraag over de helderheid van de definitie van de indicator over de ziekenhuizen. • Het vlaggetje van de betrouwbaarheid van het registratieproces, per indicator. • De verdeling van de vlaggetjes voor de betrouwbaarheid van het registratieproces over de ziekenhuizen, met het eigen ziekenhuis uitgelicht en een afspiegeling van het landelijke beeld. • Het vlaggetje van de statistische betrouwbaarheid, per indicator. • De verdeling van de vlaggetjes voor de statistische betrouwbaarheid over de ziekenhuizen, met het eigen ziekenhuis uitgelicht. • Uitspraak over de indrukvaliditeit en vergelijkbaarheid door de indicatorwerkgroep (bijlage B).
<p>Openbare database</p>	<ul style="list-style-type: none"> • Het vlaggetje van de betrouwbaarheid van het registratieproces, per indicator en de verdeling over de ziekenhuizen. • Het vlaggetje van de statistische betrouwbaarheid, per indicator en de verdeling over de ziekenhuizen. • Uitspraak over de indrukvaliditeit en vergelijkbaarheid door de indicatorwerkgroep (bijlage B). • Per indicator een leeswijzer met een algemene toelichting op het gebruik van indicatoren en voor die indicator specifieke aanwijzingen welke conclusies uit de data kunnen worden getrokken.

Bijlage D. Mogelijk hulpmiddel – definiëring indicator

Verschillende ziekenhuizen kunnen de informatie die moet worden aangeleverd verschillend interpreteren. Daardoor zijn de waarden van de indicator tussen de verschillende ziekenhuizen niet te vergelijken. Dit betekent dat de definitie van de indicator niet eenduidig is. Per indicator kan een histogram van de geobserveerde waarden worden gemaakt (zie voor een fictief voorbeeld figuur 2). Hierin kan bekeken worden of ziekenhuizen in groepen uiteen vallen. Is dit het geval, dan lijken de ziekenhuizen de aan te leveren informatie verschillend te interpreteren en is terugkoppeling naar de indicatorwerkgroep noodzakelijk. Opgemerkt dient te worden dat wanneer er in de histogrammen geen groepen te onderscheiden zijn, dit niet wil zeggen dat de definitie van de indicator eenduidig is.



Figuur 2. Histogram van proporties van een fictieve indicator.

Voor elk ziekenhuis is een proportie uitgerekend. In het histogram zijn duidelijk twee groepen ziekenhuizen te onderscheiden