



**Leeswijzer
Signaalvlag statistisch betrouwbaar
onderscheiden**

Colofon

Opdracht en eindverantwoordelijkheid

Tot stand gekomen in opdracht en onder eindverantwoordelijkheid van programmabureau Zichtbare Zorg

Auteurs

Maria Schipper, Significant

Jan Reitsma, Significant

Rosa Remmerswaal, Significant



S I G N I F I C A N T

Inhoudsopgave

| | | |
|----------|--|-----------|
| 1 | INLEIDING | 3 |
| 1.1 | Signaalvlaggen | 3 |
| 1.2 | Opbouw van het document | 4 |
| 2 | SIGNAALVLAG STATISTISCH BETROUWBAAR ONDERSCHIEDEN | 5 |
| 2.1 | Uitgangspunten signaalvlag per instelling | 5 |
| 2.2 | Benodigd aantal observaties | 7 |
| 2.3 | Uitgangspunten signaalvlag per indicator | 7 |
| 3 | ILLUSTRATIE TOEKENNEN VAN SIGNAALVLAGGEN | 8 |
| 3.1 | Toekennen signaalvlag per zorgaanbieder | 8 |
| 3.2 | Toekennen signaalvlag op indicatorniveau | 9 |
| 4 | INTERPRETATIE EN OPENBAARMAKING | 11 |
| 4.1 | Interpretatie | 11 |
| 4.2 | Openbaarmaking signaalvlaggen | 11 |
| | BIJLAGE A EXPERTBIJeenKOMST | 13 |

1 Inleiding

Bureau Zichtbare Zorg ondersteunt zorgaanbieders in verschillende sectoren bij het inzichtelijk maken van de kwaliteit van de geleverde zorg. Hiertoe ontwikkelt Zichtbare Zorg kwaliteitsindicatoren, verzamelt gegevens met betrekking tot deze indicatoren bij de zorgaanbieders en maakt deze gegevens beschikbaar voor cliënten, zorgverzekeraars, toezichthouders, zorgverleners en beleidsmakers om hen te informeren over de kwaliteit van de zorg.

1.1 Signaalvlaggen

Gebruikers van de data moeten voldoende context geboden krijgen, zodat zij verantwoord kunnen omgaan met de gegevens. Zichtbare Zorg heeft een indicatorstandaard opgesteld waarin een viertal kwaliteitseisen staat beschreven waarop indicatoren worden beoordeeld. Deze zijn als volgt gedefinieerd:

- *Inhoudsvaliditeit*. De indicator (geoperationaliseerde grootheid) meet wat hij beoogt te meten (abstract concept). De inhoudsvaliditeit op setniveau is goed als de set van indicatoren de relevante aspecten van de geleverde zorg goed dekt;
- *Vergelijkbaarheid (tussen zorgaanbieders)*. Verschillen in patiëntenpopulaties kunnen van invloed zijn op de waarde van een indicator. Wanneer zorgaanbieders met elkaar worden vergeleken, moet met deze mogelijke invloed rekening worden gehouden;
- *Registratiebetrouwbaarheid*. Het proces van meten, registreren, aanleveren en verwerken van de voor de indicator benodigde gegevens dient tijdig, juist, volledig en uniform en zonder meetfouten te gebeuren;
- *Statistisch betrouwbaar onderscheiden*. De indicator dient het vermogen te hebben om zorgaanbieders met bovengemiddelde en ondergemiddelde indicatorwaarden te onderscheiden van het landelijk gemiddelde, zonder dat toeval hier een grote rol in speelt.

De mate waarin de indicatoren voldoen aan de kwaliteitseisen wordt met behulp van signaalvlaggen inzichtelijk gemaakt.

Interpretatie

Een signaalvlag geeft aan in hoeverre een indicator aan de gestelde kwaliteitseis voldoet. De vlaggen zijn ingedeeld in drie categorieën:

1. Groen: de indicator voldoet 'voldoende' aan de gestelde kwaliteitseis;
2. Oranje: de indicator voldoet 'matig' aan de gestelde kwaliteitseis en vraagt een genuanceerd oordeel over de kwaliteit van de indicator;
3. Rood: de indicator voldoet 'onvoldoende' aan de gestelde kwaliteitseis.

De signaalvlaggen doen nadrukkelijk *geen* uitspraken over de kwaliteit van de zorg, maar vormen een indicatie voor de kwaliteit van de indicator. De signaalvlaggen zijn voor Zichtbare Zorg van belang met betrekking tot besluitvorming over het openbaar maken van de gegevens, de optimalisatie van indicatoren en de monitoring van de ontwikkeling van indicatoren over de jaren.

Signaalvlag statistisch betrouwbaar onderscheiden

In voorgaande jaren is de signaalvlag statistisch betrouwbaar onderscheiden ontwikkeld voor de sectoren Ziekenhuizen en Kraamzorg. De aldus ontstane methoden voor het berekenen van deze signaalvlag zijn enigszins verschillend. Omdat dit niet wenselijk wordt geacht, is in samenspraak met

verschillende experts (bijlage A) de signaalvlag voor statistisch betrouwbaar onderscheiden door-ontwikkeld.

Binnen Zichtbare Zorg wordt momenteel gebruikgemaakt van verschillende schattingsmethoden om indicatorwaarden te bepalen: frequentistische schattingen en Empirical Bayes-schattingen. Daarnaast zijn er indicatoren die uitgedrukt worden in proporties (bijvoorbeeld het percentage¹ cliënten met decubitus) en indicatoren die in een gemiddelde (bijvoorbeeld de gemiddelde wachttijd) uitgedrukt worden. De leeswijzer richt zich op de signaalvlag die is ontwikkeld voor indicatoren die als proportie worden uitgedrukt en met de frequentistische schattingsmethode worden geanalyseerd (zie hoofdstuk 2).

Indicatoren die betrekking hebben op een gemiddelde en indicatoren die worden geanalyseerd met Empirical Bayes-technieken, kunnen niet met de hier beschreven signaalvlag voor frequentistische analyse van proporties worden geduid. De ontwikkeling van signaalvlaggen voor gemiddelden en Empirical Bayes-methoden staat in de onderzoeksverantwoording² beschreven.

1.2 Opbouw van het document

Deze leeswijzer heeft als doel de toegekende signaalvlaggen voor statistisch betrouwbaar onderscheiden toe te lichten, zodat de gebruikers deze correct kunnen interpreteren. Voor een uitgebreide beschrijving van de totstandkoming van de signaalvlag statistisch betrouwbaar onderscheiden verwijzen wij u naar de onderzoeksverantwoording. In de volgende hoofdstukken zal worden aangegeven, wanneer nadere informatie is terug te vinden in de onderzoeksverantwoording. In deze leeswijzer zullen wij achtereenvolgens ingaan op de signaalvlaggen voor statistisch betrouwbaar onderscheiden (hoofdstuk 2), en een illustratie geven van de toekenning van de signaalvlag (hoofdstuk 3). Ter verduidelijking van de signaalvlagmethodiek zijn in blauwe tekstblokken voorbeelden gegeven.

¹ Proportie en percentage zijn twee verschijningsvormen van hetzelfde, een relatieve verhouding die ontstaat door 2 variabelen (teller en noemer) op elkaar te delen: een proportie drukt deze verhouding uit tussen 0 en 1, een percentage tussen 0% en 100%. Door een proportie met 100% te vermenigvuldigen, wordt het percentage verkregen.

² Zie *Onderzoeksverantwoording signaalvlag statistisch betrouwbaar onderscheiden* (Significant, april 2011).

2 Signaalvlag statistisch betrouwbaar onderscheiden

De kwaliteitseis statistisch betrouwbaar onderscheiden speelt alleen een rol bij proces³- en uitkomst-indicatoren. Deze indicatoren worden veelal als een teller en een noemer uitgevraagd en vervolgens uitgedrukt als een proportie.

Als indicator wordt het percentage operaties, waarbij de patiënt perioperatief geen transfusie van homolog bloed heeft gekregen in geval van een totale heupprothese, vastgesteld. De teller van deze indicator is dan gelijk aan het aantal patiënten dat perioperatief geen transfusie van homolog bloed heeft gekregen in het geval van een totale heupprothese. De noemer is gelijk aan het aantal patiënten dat een totale heupprothese heeft ondergaan.

Een indicator kan statistisch betrouwbaar onderscheiden tussen zorgaanbieders wanneer de waarden voldoende nauwkeurig zijn gemeten. Dit betekent dat toevalsfluctuaties beperkt zijn en wij vrij zeker weten dat, wanneer wij bij een nieuwe groep van cliënten dezelfde indicator opnieuw vaststellen, wij dezelfde waarde vinden⁴. De signaalvlaggen voor statistisch betrouwbaar onderscheiden worden op verschillende niveaus vastgesteld: per instelling (1.) en per indicator (2.).

1. Voor relatief kleine instellingen waarbij de indicatorwaarde op een beperkt aantal cliënten is gebaseerd speelt toeval een sterkere rol. Bij herhaalde meting zal immers de indicatorwaarde deels afhankelijk zijn van de toevallige samenstelling van het cliëntenbestand op het moment van meting. Voor grote instellingen waarbij de indicatorwaarde gebaseerd is op een groot aantal cliënten zullen toevallige verschillen ten gevolge van variërende samenstelling van het cliëntenbestand zich sneller uitmiddelen. De nauwkeurigheid van de indicatorwaarde hangt dus sterk samen met het aantal cliënten waarop deze is gebaseerd. De signaalvlag statistisch betrouwbaar onderscheiden is daarom gebaseerd op het aantal observaties waarop de indicatorwaarde is gebaseerd en wordt per zorgaanbieder vastgesteld. Wij benadrukken dat de signaalvlag *niets* zegt over de kwaliteit van zorg geleverd door een zorgaanbieder.
2. De signaalvlag voor de indicator als geheel wordt bepaald aan de hand van de signaalvlaggen voor de afzonderlijke instellingen. Ook deze signaalvlag per indicator zegt *niets* over de kwaliteit van zorg. De signaalvlag geeft aan in hoeverre instellingen over het algemeen statistisch verantwoord vergeleken kunnen worden met het landelijk gemiddelde.

2.1 Uitgangspunten signaalvlag per instelling

Het minimumaantal waarnemingen dat nodig is voor het verkrijgen van een groene of oranje signaalvlag is vastgesteld met behulp van powerberekeningen. In de statistiek worden powerberekeningen gebruikt om vast te stellen hoeveel waarnemingen nodig zijn om op een statistisch verantwoorde wijze indicatorwaarden van een groep (een zorgaanbieder) met het geheel (het landelijk gemiddelde) te vergelijken. Bij deze berekeningen spelen een drietal parameters een rol: de effectmaat, de onbetrouwbaarheid en de power. Voor deze parameters zijn uitgangspunten met experts opgesteld.

³ Er is een aantal procesindicatoren dat niet onderhevig is aan toevalsfluctuaties. Voor deze procesindicatoren speelt de kwaliteitseis statistisch betrouwbaar onderscheiden geen rol. Een voorbeeld van een dergelijke indicator is 'Voetzorg' van de indicatorset diabetes, sector Ziekenhuizen. De indicator is gedefinieerd als "Het percentage patiënten met diabetes mellitus onder behandeling van de internist, waarbij een voetonderzoek heeft plaatsgevonden".

⁴ Dit is een theoretische notie, in werkelijkheid kunnen wij niet bij een nieuwe groep patiënten dezelfde meting bij een zorgaanbieder doen.

1. De effectmaat. Om het aantal waarnemingen - dat benodigd is om statistisch verantwoord een vergelijking te maken tussen de indicatorwaarde van een zorgaanbieder en het landelijk gemiddelde – te berekenen, is het belangrijk om te bepalen in welke maat deze vergelijking wordt uitgedrukt en welke waarde van deze effectmaat als relevant wordt gezien. Bij het vergelijken van percentages liggen relatieve maten voor de hand, zeker wanneer wij te maken hebben met (zeer) kleine percentages. Hierbij zeggen wij bijvoorbeeld dat het percentage van een zorgaanbieder twee keer zo groot is als het landelijk gemiddelde percentage. Een veel gebruikte en goed gedocumenteerde effectmaat is Cohen's d, waarbij een Cohen's d van 0.2 als klein effect en een Cohen's d van 0.5 als middelmatig effect gezien. Voor percentages tussen de 10% en 90% komen beide waarden overeen met een relatief risico van ongeveer 2 en 5. Een relatief risico van 2 wil zeggen dat in een instelling de indicatorwaarde 2 keer zo hoog is al het landelijk gemiddelde, bijvoorbeeld 28% bij een landelijk gemiddelde van 14%. Voor het bepalen van de kleur van de signaalvlag worden de genoemde waarden van Cohen's d als uitgangspunt gebruikt. Voor de groene signaalvlag (goede statistische betrouwbaarheid) geldt een klein effect als criterium (Cohen's d = 0.2) en voor een oranje signaalvlag (matige statistische betrouwbaarheid) een middelmatig effect (Cohen's d = 0.5). Is ook een middelmatig effect statistisch onvoldoende betrouwbaar vast te stellen, dan wordt een rode vlag gegeven.⁵
2. Er wordt bij de berekening van het aantal observaties uitgegaan van een significantieniveau⁶ van 5% en een onderscheidingsvermogen (power⁷) van 80%. Er is hierbij geen nadere uitsplitsing naar de kleur van de vlag. De gehanteerde grenzen zijn welliswaar gangbaar, maar redelijk liberaal. De hieruit berekende aantallen observaties moeten hierdoor als ondergrenzen worden geïnterpreteerd.

Naast de effectmaat, het significantieniveau en onderscheidend vermogen zijn er nog twee uitgangspunten die met name gaan om het detailniveau: de vorm van statistische toetsing⁸ die wordt gebruikt, en het aantal categorieën waarmee wordt gerekend:

1. Het landelijk gemiddelde (gewogen naar het aantal observaties per instelling) van de indicator wordt gezien als de norm die in de statistische toetsing wordt gebruikt;
2. De landelijk gemiddelde proportie waarmee de proporties van instellingen worden vergeleken, is in vijf categorieën ingedeeld (zie hiervoor paragraaf 2.2).

⁵ Voor proporties waarbij het landelijk gemiddelde erg dicht bij 0 of 1 ligt, wordt de Cohen's d vertaald naar een odds ratio. Zie ook de onderzoeksverantwoording.

⁶ Het significantieniveau: de kans dat een statistisch gevonden verschil op toeval berust en er in werkelijkheid geen sprake is van verschil.

⁷ Onderscheidend vermogen: de kans dat een in werkelijkheid bestaand verschil ook door middel van statistische analyse gevonden kan worden.

⁸ Statistische toets: in dit geval een vergelijking of de indicatorwaarde van de instelling statistisch afwijkt van de landelijk gemiddelde proportie.

2.2 Benodigd aantal observaties

Aan de hand van bovenstaande uitgangspunten is voor de verschillende kleuren vlaggen het minimumaantal benodigde observaties per zorgaanbieder berekend.

| Landelijk gemiddelde (p) | | Minimaal aantal observaties | |
|--------------------------|----------------------|-----------------------------|-------------|
| | | Groene vlag | Oranje vlag |
| $p \leq 0,02$ | $p \geq 0,98$ | 2300 | 640 |
| $0,02 < p \leq 0,04$ | $0,98 > p \geq 0,96$ | 1160 | 320 |
| $0,04 < p \leq 0,07$ | $0,96 > p \geq 0,93$ | 580 | 160 |
| $0,07 < p \leq 0,12$ | $0,93 > p \geq 0,88$ | 340 | 90 |
| $0,12 < p < 0,88$ | | 200 | 30 |

Tabel 1: Minimumaantallen observaties per zorgaanbieder voor proporties

In tabel 1 is weergegeven wat de minimale benodigde observaties per zorgaanbieder zijn, afhankelijk van het landelijk gemiddelde, voor de groene en oranje vlag. Een rode vlag wordt aan de zorgaanbieder toegekend, indien niet kan worden voldaan aan de eisen van de oranje vlag.

2.3 Uitgangspunten signaalvlag per indicator

De beoordeling van de statistische betrouwbaarheid voor de indicator - voor alle zorgaanbieders gezamenlijk - wordt gebaseerd op de signaalvlaggen van alle zorgaanbieders:

1. Wanneer 25% of meer van de instellingen een rode signaalvlag heeft, dan is de statistische betrouwbaarheid voor de betreffende indicator 'slecht' en krijgt deze indicator als geheel ook een rode signaalvlag;
2. Indien 75% of meer van de instellingen een groene signaalvlag heeft, dan krijgt de indicator een groene signaalvlag en is de statistische betrouwbaarheid 'goed';
3. Als minder dan 25% van de instellingen rood en minder dan 75% van de instellingen groen is, dan krijgt de indicator een oranje signaalvlag en is de statistische betrouwbaarheid 'matig'.

Aan de hand van de signaalvlaggen per indicator, worden de volgende adviezen meegegeven aan de gebruikers:

| | |
|--------|--|
| Rood | Alle zorgaanbieders bezien, is een zinnige vergelijking van indicatorwaarden van zorgaanbieders met het landelijke gemiddelde niet mogelijk. |
| Oranje | Alle zorgaanbieders bezien, vraagt een zinnige vergelijking van indicatorwaarden van zorgaanbieders met het landelijk gemiddelde een genuanceerd oordeel. |
| Groen | Alle zorgaanbieders bezien, is een zinnige vergelijking van indicatorwaarden van zorgaanbieders met het landelijke gemiddelde statistisch verantwoord. Echter er kunnen nog verschillen in cliëntkenmerken (zogenaamde case mix-variabelen) zijn die verstoring werken (zie kwaliteitsdimensie vergelijkbaarheid). |

3 Illustratie toekennen van signaalvlaggen

In hoofdstuk 2 zijn de uitgangspunten en methoden besproken voor het toekennen van de signaalvlag statistisch betrouwbaar onderscheiden. In dit hoofdstuk worden voorbeelden besproken ter verduidelijking van de wijze van toekenning van de signaalvlaggen.

In tabel 2 is voor tien fictieve zorgaanbieders (a t/m j) een voorbeeld opgenomen voor het toekennen van de signaalvlag voor twee verschillende indicatoren. Per zorgaanbieder zijn de proporties en de noemers (het aantal observaties) weergegeven. Aan de hand van deze tabel zal in dit hoofdstuk de toekenning van signaalvlaggen worden geïllustreerd.

| | indicator 1 | | | indicator 2 | | |
|-------------------|-------------|--------------------|-------------|-------------|--------------------|-------------|
| | Proportie | Aantal observaties | Signaalvlag | Proportie | Aantal observaties | Signaalvlag |
| Zorgaanbieder a | 0,06 | 465 | | 0,88 | 396 | |
| Zorgaanbieder b | 0,12 | 264 | | 0,94 | 258 | |
| Zorgaanbieder c | 0,04 | 987 | | 0,93 | 694 | |
| Zorgaanbieder d | 0,05 | 759 | | 0,84 | 582 | |
| Zorgaanbieder e | 0,04 | 148 | | 0,87 | 91 | |
| Zorgaanbieder f | 0,11 | 198 | | 0,73 | 214 | |
| Zorgaanbieder g | 0,03 | 396 | | 0,78 | 352 | |
| Zorgaanbieder h | 0,06 | 118 | | 0,85 | 83 | |
| Zorgaanbieder i | 0,05 | 583 | | 0,82 | 418 | |
| Zorgaanbieder j | 0,08 | 269 | | 0,87 | 279 | |
| Totaal | 0,064 | | | 0,851 | | |
| Grens oranje vlag | 160 | | | 30 | | |
| Grens groene vlag | 580 | | | 200 | | |

Tabel 2: Voorbeeld toekenning signaalvlaggen frequentistische methode.

3.1 Toekennen signaalvlag per zorgaanbieder

Voor het bepalen van de grenzen van de verschillende kleuren voor de signaalvlaggen, dient eerst het (naar aantal observaties gewogen) landelijk gemiddelde te worden berekend. Voor indicator 1 en 2 is deze te vinden in de rij 'totaal' in de kolom 'proportie'. Met behulp van tabel 1 zijn de grenzen voor het minimumaantal benodigde observaties bepaald en onderaan in tabel 2 weergegeven. Per zorgaanbieder is de kleur van de signaalvlag bepaald aan de hand van deze grenzen. De kleur van de signaalvlag zegt *niets* over de kwaliteit van de zorg, maar alleen of de indicatorwaarde op voldoende observaties is gebaseerd, om een statistisch betrouwbare vergelijking te maken met het landelijk gemiddelde.

Voor twee aanbieders (zorgaanbieder a en e) is het toekennen van de signaalvlag nader uitgewerkt in de blauwe tekstvakken.

Zorgaanbieder a

De proportiescore voor indicator 1 is gebaseerd op 465 waarnemingen. Dit aantal ligt boven de grens van de oranje vlag (minimaal 160 waarnemingen), maar nog onder het benodigde aantal waarnemingen voor een groene vlag (minimaal 580 waarnemingen). Om deze reden is een oranje vlag toegekend. De statistische betrouwbaarheid wordt dan als 'matig' gekwalificeerd, dit houdt in dat wanneer een zorgaanbieder significant afwijkt van het gemiddelde, er een gematigde kans is dat dit gebaseerd is op toeval. De vergelijking van deze zorgaanbieder met het landelijk gemiddelde vraagt om een genuanceerd oordeel.

De proportiescore voor indicator 2 is gebaseerd op 396 waarnemingen. Dit aantal ligt boven de grens van de groene vlag (minimaal 200 waarnemingen), zodat er een groene vlag wordt toegekend. De statistische betrouwbaarheid wordt 'goed' genoemd, wat inhoudt dat deze zorgaanbieder zich statistisch betrouwbaar kan onderscheiden van het landelijk gemiddelde.

Zorgaanbieder e

De proportiescore voor indicator 1 is gebaseerd op 148 waarnemingen. Dit aantal ligt onder de grens van de oranje vlag (minimaal 160 waarnemingen). Om deze reden is een rode vlag toegekend; de statistische betrouwbaarheid wordt 'slecht' genoemd. De kans is daarom groot dat een statistisch significante afwijking berust op toeval. Het is daarom niet verantwoord deze vergelijking te maken.

De proportiescore voor indicator 2 is gebaseerd op 91 waarnemingen. Dit aantal ligt boven de grens van de oranje vlag (minimaal 30 waarnemingen), maar nog onder de grens voor een groene vlag (minimaal 200 waarnemingen). Derhalve is een oranje vlag toegekend; de statistische betrouwbaarheid wordt 'matig' genoemd. De vergelijking van deze zorgaanbieder met het landelijk gemiddelde vraagt om een genuanceerd oordeel.

3.2 Toekennen signaalvlag op indicatorniveau

In paragraaf 2.2 zijn de grenzen opgenomen voor de signaalvlag op indicatorniveau. In de onderstaande tekstvakken wordt voor de indicatoren uit tabel 2 toegelicht hoe de signaalvlag voor de indicator wordt bepaald.

Indicator 1:

- Rode vlag: 20% van de zorgaanbieders;
- Oranje vlag: 50% van de zorgaanbieders;
- Groene vlag: 30% van de zorgaanbieders.

Aan minder dan 75% van de zorgaanbieders is een rode vlag toegekend; tevens zijn er aan minder dan 75% van de zorgaanbieders een groene vlag toegekend. Aangezien aan geen van beide grenzen wordt voldaan, is een oranje vlag toegekend. Alle zorgaanbieders bezien, vraagt een zinnige vergelijking van zorgaanbieders met het landelijke gemiddelde om een genuanceerd oordeel.

Indicator 2:

- Rode vlag: 0% van de zorgaanbieders;
- Oranje vlag: 20% van de zorgaanbieders;
- Groene vlag: 80% van de zorgaanbieders.

Aan meer dan 75% van de zorgaanbieders is een groene vlag toegekend. Om deze reden is voor de statistische betrouwbaarheid van de indicator ook een groene vlag toegekend. Alle zorgaanbieders bezien, is een zinnige vergelijking van zorgaanbieders met het landelijke gemiddelde op statistisch verantwoord. Echter er kunnen nog case mix-variabelen zijn die verstorend werken (zie populatie-vergelijkbaarheid).

4 Interpretatie en openbaarmaking

4.1 Interpretatie

De signaalvlaggen moeten worden beschouwd als niet meer maar ook niet minder dan 'richtlijn'. Voor statistisch betrouwbaar onderscheiden komt dit door de volgende factoren:

1. De grove indeling in drie verschillende kleuren;
2. De afhankelijkheid van interpretatie van de signaalvlag van de specifieke doelen van de informatiegebruiker. Een cliënt gebruikt de informatie bijvoorbeeld met het doel een keuze te maken tussen verschillende zorgaanbieders. Voor een verzekeraar speelt naast kwaliteit ook prijs een rol. Toezichthouders gebruiken de indicatoren - onder meer - als voorspellers van onveilige situaties. Het hangt van deze doelen af of men geïnteresseerd is in kleine verschillen in indicatorwaarden tussen de zorgaanbieders of juist grote verschillen. De signaalvlag is onder andere gebaseerd op een aantal vaste (absolute) verschillen. De toegewezen signaalvlag kan dan door gebruikers als geschikt (ze zijn in dezelfde verschillen als gesteld in de signaalvlag geïnteresseerd), te conservatief (ze vinden grotere verschillen interessant) of liberaal (juist kleinere verschillen zijn interessant) worden geïnterpreteerd.⁹

Uit tabel 2 blijkt dat zorgaanbieder f voor de eerste indicator een oranje vlag heeft toegewezen gekregen. Bij de berekening van de minimumaantal benodigde observaties is ervan uitgegaan dat de gebruiker van de data geïnteresseerd is in een relatief risico, gelijk aan 2. Indien een gebruiker echter geïnteresseerd is in grotere absolute verschillen, bijvoorbeeld een relatief risico gelijk aan 4, zou de proportie van de zorgaanbieder gebaseerd zijn op voldoende observaties. In dit geval is de toegewezen vlag te conservatief.

4.2 Openbaarmaking signaalvlaggen

Voor iedere indicator wordt een signaalvlag voor elk van de vier kwaliteitseisen vastgesteld. Op basis van deze signaalvlaggen kan de stuurgroep¹⁰ van Zichtbare Zorg een onderbouwd besluit nemen over het al dan niet (of deels) openbaar maken van de data. Het uitgangspunt daarbij is dat alle indicatorwaarden openbaar worden gemaakt, samen met de signaalvlaggen op indicatorniveau. Iedere zorgaanbieder krijgt voor de kwaliteitseisen 'registratiebetrouwbaarheid' en 'statistisch betrouwbaar onderscheiden' een individuele signaalvlag toegekend. De stuurgroep kan echter besluiten om deze individuele signaalvlaggen niet vrij te geven. Een reden hiervoor kan zijn dat het risico bestaat dat de signaalvlag zelf als kwaliteitsinformatie wordt gezien in plaats van de indicatorwaarde.

⁹ Naast relatieve verschillen kan een gebruiker ook geïnteresseerd zijn in absolute verschillen. Gezien de signaalvlag is ontwikkeld om relatieve verschillen te signaleren, zal ook voor deze gebruiker de signaalvlag niet altijd bruikbaar zijn.

¹⁰ Elke sector vertegenwoordigd bij Zichtbare Zorg heeft een stuurgroep. Deze stuurgroep is eindverantwoordelijk voor de samenstelling en meting van indicatorensets in portefeuille en neemt daarvoor de belangrijke, strategische beslissingen. Vertegenwoordigers van zorgaanbieders, brancheorganisaties, cliëntenorganisaties, verzekeraars en de IGZ hebben zitting in de stuurgroep.

Een ziekenhuis heeft voor de indicator 'Percentage operaties waarbij de patiënt perioperatief geen transfusie van homolog bloed heeft gekregen in geval van een totale heupprothese' als indicatorwaarde 99%. De signaalvlag voor statistisch betrouwbaar onderscheiden is voor dit ziekenhuis rood. Gebruikers van de data zouden deze vlag kunnen interpreteren als *dit ziekenhuis scoort slecht op deze indicator*. Deze interpretatie is niet correct: het ziekenhuis scoort goed op deze indicator, het is alleen niet statistisch verantwoord deze waarde met de landelijk gemiddelde score te vergelijken.

Bijlage A Expertbijeenkomst

De signaalvlag statistisch betrouwbaar onderscheiden is in samenspraak met verschillende experts ontwikkeld. Tijdens twee expertbijeenkomsten in de periode februari - maart 2011 zijn de verschillende uitgangspunten besproken.

De volgende experts waren aanwezig bij deze bijeenkomsten.

| Achternaam | Voorletters | Voornaam | Titelatuur | Organisatie | Afdeling |
|---------------|-------------|----------|------------|----------------------------|--|
| Anema | | Helen | | AMC | public health |
| Beek, van | | Sandra | | Actiz | |
| Berg, van den | M.J. | Michael | dr. | RIVM | Zorgbalans |
| Boer, de | D. | Dolf | dr. | Centrum Klantervaring Zorg | |
| Braspenning | | Jozé | | UMC St Radboud | IQ Healthcare |
| Koolman | A.H.E. | Xander | dr. | TU Delft/IPSE studies | Universitair hoofddocent Publieke sector efficiëntie |
| Rij | C. | Coen | dr. | Regioplan | Clustermanager Arbeid en economie |
| Smits | I. | Ilja | | Vektis | |
| Steyerberg | E.W. | Ewout | prof. dr. | Erasmus MC | Maatschappelijke Gezondheidszorg |
| Tempelman | C. | Caren | dr. | SEO | |
| Zwinderman | A.H. | Koos | prof. dr. | AMC | Klinische epidemiologie en biostatistiek |