

Postprint Version	1.0
Journal website	http://vb23.bsl.nl/frontend/redirect.asp?page=1388-7491/tsg86-447.pdf
Pubmed link	
DOI	

This is a NIVEL certified Post Print, more info at <http://www.nivel.eu>

De constructie van een CQI meetinstrument: ervaringen uit de praktijk.

JANY RADEMAKERS* , HERMAN SIXMA* , MATTANJA TRIEMSTRA* , OLGA DAMMAN* , MICHELLE HENDRIKS* , MARLOES ZUIDGEEST*

* NIVEL, Nederlands instituut voor onderzoek van de gezondheidszorg, Utrecht
Jany Rademakers, NIVEL, Postbus 1568, 3500 BN Utrecht, tel. 030-2729847, e-mail: j.rademakers@nivel.nl

THE CONSTRUCTION OF A CQ-INDEX QUESTIONNAIRE: PRACTISE EXPERIENCES

The development of a CQ-index questionnaire consists of several phases. In this article the construction of a questionnaire and the statistical analyses with regard to the psychometrical properties and discriminative capacity of the instrument are discussed. The description will be illustrated with examples from four CQI's: the CQI General Practice, the CQI Hip/knee-surgery, the CQI Mammacare and the CQI Reumatoid Arthritis. The experiences with different phases of the construction process as well as with the CQI systematics in general will be discussed.

De ontwikkeling van een CQI meetinstrument bestaat uit verschillende fasen. In dit artikel wordt ingegaan op de constructie van de vragenlijst en op de statistische analyses ten aanzien van de psychometrische aspecten en het discriminerend vermogen. De beschrijving van deze fasen zal geïllustreerd worden aan de hand van voorbeelden uit de CQI Huisartsenzorg Overdag, de CQI Heup/knie-operatie, de CQI Mammacare en de CQI Reumatoïde Arthritis. Verder zullen de ervaringen met verschillende onderdelen van het constructieproces en de CQI meetsystematiek als geheel worden bediscussieerd.

INLEIDING

De Consumer Quality- index (CQ-index of CQI) is een combinatie van twee families van vragenlijsten die ontwikkeld zijn voor het meten van kwaliteit van zorg vanuit het perspectief van patiënten: de Amerikaanse CAHPS (Consumer Assessment of Healthcare Providers and Systems) (<http://www.cahps-sun.org/home/index.asp>) en de Nederlandse QUOTE (Quality Of care Through the patient's Eyes), ontwikkeld door het NIVEL.¹⁻⁴

Wat beide families gemeenschappelijk hebben is dat ze ervaringen meten van consumenten van zorg. Dat is een belangrijke verbetering ten opzichte van vragenlijsten uit de jaren '80, waarin alleen naar het meer subjectieve begrip tevredenheid werd gevraagd.⁵ Beide meetsystemen zijn wetenschappelijk gevalideerd en getest. De reden om de twee families te combineren tot één nieuwe meetsystematiek was dat de CAHPS weliswaar de meest objectieve antwoordsystematiek en gestandaardiseerde procedures rondom

steekproeftrekking, dataverzameling, analyse en rapportage bood, maar dat de QUOTE het belang dat mensen hechten aan specifieke kwaliteitsaspecten meenam en de mogelijkheid had om vragenlijsten voor specifieke aandoeningen, patiëntenpopulaties of medische interventies te ontwikkelen.

Een CQ-index heeft dus de volgende eigenschappen:

Er wordt zowel bij het ontwikkeltraject als bij reguliere metingen gebruik gemaakt van gestandaardiseerde procedures, om bias bij de steekproeftrekking, dataverzameling, analyse en rapportage te voorkomen en optimale vergelijkbaarheid van de gegevens te bevorderen.

De lijst kan betrekking hebben op een sector, groep aanbieders, aandoening of medische ingreep, afhankelijk van het doel waarvoor de lijst ontwikkeld wordt.

Er wordt bij de ontwikkeling van het instrument en periodiek bij reguliere metingen gevraagd naar het belang dat mensen hechten aan bepaalde kwaliteitsaspecten en bij alle metingen naar de ervaring die ze er mee hebben.

De vraag- en antwoordformuleringen hebben een specifiek format waarvan in principe niet mag worden afgeweken. Daarnaast is er een aantal verplichte vragen dat in alle lijsten moet worden opgenomen.

Doel van deze standaardisatie is dat de gegevens die met de CQ-index worden verzameld betrouwbaar, valide en vergelijkbaar zijn. Tegelijkertijd leidt standaardisatie tot beperking van de mogelijke vraagstellingen en antwoordcategorieën. Door de set van verplichte vragen worden de lijsten langer. Door beide punten kunnen de validiteit en de betrouwbaarheid in het geding komen. Maatwerk is wel weer mogelijk door instrumenten voor specifieke aandoeningen of ingrepen te ontwikkelen, maar dat betekent ook dat er veel verschillende lijsten moeten worden geconstrueerd. Hierbij spelen kosten en belangen een rol.

In dit artikel zullen we – aan de hand van enkele voorbeelden – het constructieproces van een CQI meetinstrument beschrijven. Daarbij wordt de beschrijving van de standaardprocedures en richtlijnen aangevuld met ervaringen uit de praktijk. In de discussie zullen we nader ingaan op de eerder genoemde thema's rondom standaardisatie en maatwerk.

HET CONSTRUCTIEPROCES VAN EEN CQI MEETINSTRUMENT

Het constructieproces van een CQI meetinstrument dient te gebeuren volgens de procedures en richtlijnen die staan beschreven in het Handboek.⁶

De volgende fasen maken deel uit van de ontwikkeling van een CQ-index (figuur 1):

- Voorbereidende fasen (1 tot en met 4)
- Constructiefase (fase 5)
- Literatuurstudie / deskresearch
- Kwalitatief onderzoek (focusgroepen met patiënten/cliënten)
- Opstellen en uittesten eerste versie CQ-index meetinstrument
- Dataverzameling (fase 6)
- Statistische analyses (fase 7)

Test van de psychometrische kwaliteiten, w.o. het vaststellen van schalen die op een betrouwbare manier kwaliteitsdimensies kunnen meten

Test van het discriminerend vermogen waarin wordt nagegaan op welke kwaliteitsdimensies statistisch significante verschillen tussen instellingen of aanbieders kunnen worden vastgesteld

- Rapportage over de voorafgaande fasen met aanbevelingen voor de definitieve versie van de CQ-index vragenlijst (fase 8)
- Nieuwe versie CQI vragenlijst (fase 9) en vaststellen CQI vragenlijst

In de meeste CQI ontwikkeltrajecten worden de test van de psychometrische aspecten en de test van het discriminerend vermogen in twee aparte –elkaar opvolgende– projecten uitgevoerd met elk hun eigen dataverzameling. Belangrijk voordeel hiervan is dat tussentijds de vragenlijst geoptimaliseerd kan worden voordat in een grotere steekproef het

discriminerend vermogen wordt vastgesteld. In enkele gevallen gebeuren deze tests tegelijkertijd, op basis van de gegevens van één dataverzameling.

In dit artikel wordt met name ingegaan op fase 5 en 7 van het ontwikkeltraject; het construeren en empirisch toetsen van een CQI vragenlijst. Hierbij wordt gebruik gemaakt van voorbeelden en ervaringen uit diverse projecten. Hierbij is gekozen voor:

- CQI Huisartsenzorg Overdag⁷
- CQI Heup/knie-operatie^{8,9}
- CQI Mammacare^{10,11}
- CQI Reumatoïde Artritis^{12,13}

Deze lijsten zijn gekozen omdat zij verschillende soorten CQ-indexen betreffen (per groep aanbieders, ingreep en aandoening) en het hele ontwikkeltraject hebben doorlopen.

Verschiede manieren van dataverzameling (fase 6) worden elders in dit themanummer beschreven.¹⁴

[FIGUUR 1]

CONSTRUCTIEFASE

Literatuuronderzoek / deskresearch / werkgroep van 'stakeholders'

Aan het begin van elk CQI ontwikkeltraject wordt gestart met literatuuronderzoek: de onderzoeker gaat na wat er aan richtlijnen en/of bestaande meetinstrumenten voorhanden is en wat er eerder aan klantveringsonderzoek bij deze doelgroep of in deze sector is verricht. Ook wordt gekeken naar kenmerken van de doelgroep, de aandoening of de sector. Het betreft echter geen systematische review volgens de wetenschappelijke richtlijnen.

In dezelfde periode wordt een werkgroep opgericht waarin naast de onderzoekers en een vertegenwoordiger van het CKZ alle 'stakeholders' vertegenwoordigd zijn (patiënten of patiëntvertegenwoordigers, zorgaanbieders, zorgverzekeraars of andere zorginkopers zoals gemeenten). Zo mogelijk worden bij de samenstelling van de werkgroep aanbieders betrokken die bij de richtlijnontwikkeling actief waren. In de eerste bijeenkomst van deze werkgroep wordt een nauwkeurige afbakening gemaakt van het domein en de doelgroep waar de CQI betrekking op zal hebben. Dit gebeurt aan de hand van kenmerken van de aandoening (soms geoperationaliseerd in bepaalde DBC-codes), van het zorgproces, van kenmerken van de organisatie van het zorgaanbod of kenmerken van de doelgroep. Ook wordt er besloten welke onderdelen van het zorgproces of welke zorgverleners wel en niet in aanmerking komen om opgenomen te worden in de betreffende CQ-index. Deze keuzes worden gemaakt op grond van verschillende criteria, bijvoorbeeld dat er voor andere delen al een aparte lijst bestaat of dat een betreffende zorgverlener niet standaard in het proces of de keten is opgenomen. In de werkgroep Mammacare is bijvoorbeeld de keuze gemaakt om alle patiënten met een borstafwijking als doelgroep te kiezen en niet uitsluitend diegenen met de diagnose borstkanker. Bij huisartsenzorg werd expliciet gekozen voor de zorg overdag, omdat in een separaat traject al een CQ-index voor de Huisartsenposten werd ontwikkeld.

Kwalitatief onderzoek

Omdat de CQ-index tot doel heeft de kwaliteit van zorg vanuit het perspectief van patiënten en cliënten te meten is hun inbreng bij het tot stand komen van de vragenlijst essentieel. Deze inbreng wordt gewaarborgd door een kwalitatieve fase aan het begin van het project. Bij de meeste projecten zijn twee of drie focusgroepen gehouden met een beperkt aantal patiënten: CQI Huisartsenzorg (N=14), CQI Heup/knie-operatie (N= 31), CQI Mammacare (N=27) en CQI Reumatoïde Artritis (N=22). Overigens kunnen ook andere kwalitatieve methoden worden gebruikt als dat beter aansluit bij de doelgroep (interviews, internetforum). De kwalitatieve fase heeft tot doel de validiteit van de vragenlijst te waarborgen.

Zoals gebruikelijk bij focusgroepsdiscussies¹⁵ hebben de groeps gesprekken in de regel een open, exploratief karakter. Er wordt begonnen met een openingsvraag, waarop de

deelnemers vrij kunnen reageren. De gesprekken worden opgenomen op band, integraal uitgetypt en gecodeerd, waar mogelijk door twee personen. Daarbij wordt gekeken welke kwaliteitsaspecten en thema's in de gesprekken genoemd werden en hoe vaak.

In alle projecten blijken de focusgroepen een goede bron van informatie over wat patiënten belangrijke kwaliteitsaspecten vinden. Als we de thema's die genoemd worden in de focusgroepen vergelijken met wat (later) uit de gegevens van de betreffende CQI belangvragenlijsten naar voor komt dan is er een grote mate van overlap in wat de deelnemers van focusgroepen belangrijk vinden in de zorg en datgene wat later in de vragenlijsten als prioriteit wordt aangemerkt.¹⁶

Opstellen en uittesten conceptvragenlijst

Op basis van de resultaten van de kwalitatieve fase wordt een conceptvragenlijst opgesteld. Elke CQ-index bestaat uit twee afzonderlijke lijsten: de ervaringsvragenlijst en de belangvragenlijst. De ervaringsvragenlijst meet de ervaringen met de zorg in een afgebakende periode en bestaat uit een aantal samenhangende sets van vragen. Bij deze vragen wordt onderscheid gemaakt tussen:

1. Verplichte achtergrondvragen (bijvoorbeeld sociaal-demografische kenmerken)
2. Facultatieve achtergrondvragen (bijvoorbeeld over de aandoening of de kwaliteit van leven van de respondent)
3. Verplichte vragen over de algemene kwaliteit van zorg, op een standaard manier geformuleerd (voorbeelden: vragen naar de bejegening van cliënten door zorgverleners, vragen naar communicatie en informatie door de zorgverleners; algemene waarderingsvragen of 'rapportcijfers')
4. Specifieke vragen over de kwaliteit van zorg, toegesneden op de doelgroep van het nieuw te ontwikkelen CQI meetinstrument (voorbeelden: vragen over specifieke onderzoeken of behandelingen, vragen over nazorg)
5. Selectie- of screenervragen, bedoeld om de respondent op een zo efficiënt mogelijke manier door de vragenlijst te leiden.

Een aantal vragen moet verplicht worden opgenomen in de vragenlijst. Dit is onder meer om vergelijking tussen CQI instrumenten en doelgroepen mogelijk te maken en om verschillen in de populaties van verschillende instellingen vast te kunnen stellen zodat daar later op gecorrigeerd kan worden. Daarnaast moeten de vragen en antwoordcategorieën op een gestandaardiseerde manier geformuleerd worden:

- Frequentie waarmee iets gebeurt is (nooit – soms – meestal – altijd).
- Mate waarin iets een probleem was (groot – klein – geen probleem)
- Selectievragen of andere dichotome vragen (ja – nee)
- Algemene waarderingsvragen (schaal van 0 – 10)

De belangvragenlijst wordt uit de ervaringsvragenlijst afgeleid en heeft tot doel vast te stellen hoe belangrijk patiënten en cliënten de afzonderlijk kwaliteitsaspecten van de zorg vinden (niet belangrijk – eigenlijk wel belangrijk – belangrijk – van het aller grootste belang). Op basis van de combinatie van de gegevens uit beide vragenlijsten kan berekend worden waar verbetermogelijkheden zijn met betrekking tot de zorgkwaliteit en welke, voor de patiënt belangrijke aspecten, juist goed worden beoordeeld. De belangvragenlijst wordt bij alle ontwikkeltrajecten afgenomen en is dus een integraal onderdeel van het constructieproces. Op basis van de belangscores kan mede beslist worden of sommige items al dan niet moeten vervallen uit de definitieve versie van de vragenlijst. Bij reguliere metingen hoeft een belanglijst niet elke keer te worden afgenomen, maar kunnen de bestaande belangscores meerdere malen gebruikt worden. De aanname is dat deze scores redelijk stabiel zijn over de tijd.

De conceptvragenlijsten – die psychometrisch getest gaan worden – variëren in lengte. Vragenlijsten die betrekking hebben op een chronische ziekte of een zorgproces waarbij verschillende hulpverleners betrokken zijn, zijn gemiddeld langer dan vragenlijsten die betrekking hebben op een operatie (zonder voortraject of nazorg) of een contact met één

hulpverlener. De conceptvragenlijst van de CQI Huisartsenzorg overdag had 81 items, de CQI Heup/knie-operatie 93, de CQI Reumatoïde Artritis 142 en de (oorspronkelijke) CQI Mammacare 166. Soms wordt de lengte van de vragenlijsten naar voor gebracht als een mogelijk nadeel van het instrument. De verwachting daarbij is dat een lange vragenlijst een negatief effect heeft op de respons. Uit de respons in de (eerste) testfase blijkt er echter geen omgekeerd evenredige relatie tussen de lengte van een lijst en de mate waarin dit een belemmering voor respondenten is om de lijst in te vullen en terug te sturen: Bij de CQI Huisartsenzorg, de kortste lijst van onze voorbeelden, was de bruto-respons het laagst (61%), bij de CQI Mammacare, de langste, was de respons 63%. De andere twee lijsten die daar wat betreft het aantal items tussen zaten hadden beide zeer hoge responspercentages: CQI Heup/knie-operatie 77%, CQI Reumatoïde Artritis 74%. De respons op een vragenlijst om klantervaringen te meten wordt kennelijk door andere factoren dan de lengte bepaald.

Over de vraagformuleringen en antwoordcategorieën is al sinds de invoering van de CQI systematiek enige discussie. Eén thema is de begrijpelijkheid voor respondenten. Naar aanleiding van onderzoek van een in taal gespecialiseerd bureau is de set van verplichte basisvragen aangepast naar een eenvoudiger niveau. De Universiteit van Amsterdam, afdeling sociale geneeskunde, doet in samenwerking met het NIVEL onderzoek naar de begrijpelijkheid en validiteit van de vragen bij allochtone doelgroepen. Met betrekking tot de antwoordcategorieën doen zich twee hoofdproblemen voor. Ten eerste is het bij eenmalige contacten niet mogelijk te vragen naar frequentie en moet daarom gekozen worden voor ja/nee vragen, die moeilijker schaalbaar zijn. Ten tweede zijn er vragen waar men behoefte heeft aan meer nuancering in het antwoord dan een tweedeling. Dit geldt bijvoorbeeld bij vragen over informatievoorziening: de vraag is niet alleen óf er informatie verstrekt is maar of die informatie voldoende en kwalitatief goed was. Om deze redenen komen incidenteel in CQI vragenlijsten andere typen vragen voor, zoals in de CQI Reumatoïde Artritis: 'Heeft u goede voorlichting gekregen wat betreft de schadelijkheid van medicijnen voor vruchtbaarheid en/of het ongeboren kind en/of het verloop van de aandoening tijdens de zwangerschap? Ja, voldoende; ja, maar onvoldoende, nee'. Aangezien de term 'voldoende' afhankelijk is van een subjectieve maatstaf sluit deze vorm van vragen eigenlijk niet aan bij de uitgangspunten van de CQ-index.

STATISTISCHE ANALYSES

Psychometrische test van de conceptvragenlijst

Nadat de conceptvragenlijst is vastgesteld volgt de fase van steekproeftrekking en dataverzameling.¹⁴ Als de vragenlijsten ontvangen zijn en de data zijn ingevoerd wordt de dataset geschoond volgens de richtlijnen in het Handboek. De opgeschoonde databestanden vormen het uitgangspunt voor de statistische analyses, die dienen om de psychometrische kwaliteit van de vragenlijst vast te stellen. Daarbij wordt gekeken naar: a) de kwaliteit en bruikbaarheid van afzonderlijke vragen (items); b) de opbouw van het meetinstrument in kwaliteitsdimensies en/of thema's (schalen); c) de stabiliteit en betrouwbaarheid van deze schalen; en d) de relevantie van de kwaliteitsaspecten.

Om vast te stellen of een afzonderlijk item verwijderd kan of moet worden uit de vragenlijst worden de volgende analyses uitgevoerd:

- Itemanalyses (ervaringsvragen en belangvragen). Extreem verdeelde items (>90% van de antwoorden in extreme categorie), items met een hoge non-respons (>5% of >10% missende waarden, afhankelijk van de kwaliteit van de data; bij de CQI Huisartsenzorg en Heup/knie-operatie is uitgegaan van >10%) en items met een zeer lage gemiddelde belangscore komen in principe in aanmerking om verwijderd te worden.
- Inter-item analyses (ervaringsvragen). Items die een sterke onderlinge samenhang vertonen (Pearson's correlatie van $r > 0,70$) komen in aanmerking om verwijderd te worden (één van de twee items zou verwijderd kunnen worden).

- Factoranalyses (ervaringsvragen). Items die niet passen binnen bepaalde thema's of dimensies (factorlading van $<0,40$) komen in aanmerking om verwijderd te worden.
- Betrouwbaarheidsanalyses (ervaringsvragen). Items die niet bijdragen aan de betrouwbaarheid van een schaal (Cronbach's alpha $<0,70$; ITC (Item Totaal Correlatie) $<0,40$) komen in aanmerking om verwijderd te worden.

Deze criteria wegen allen in gelijke mate mee bij voorstellen om items al dan niet te verwijderen. Items waar meerdere criteria op van toepassing zijn hebben een grotere kans om uit de lijst geschrapt te worden. In de CQI Huisartsenzorg was één vraag met een non-respons boven de 10% en twee met een extreme verdeling. Bij de CQI Heup/knie-operatie was op geen enkel item een non-respons boven de 10%, wel hadden 13 items een extreme verdeling. Bij de CQI Mammacare hadden 43 items meer dan 5% missende waarden. In totaal waren 15 items extreem verdeeld. Bij de CQI Reumatoïde Artritis hadden 15 items meer dan 5% missende waarden en waren 13 items extreem verdeeld.

Of een item ook daadwerkelijk verwijderd wordt is afhankelijk van een aantal factoren, zoals de mate waarin de resultaten van de verschillende analyses elkaar bevestigen of tegenspreken, en het belang dat aan het kwaliteitsaspect wordt toegekend (door de patiënt/cliënt of de werkgroep).

De onderliggende schaalstructuur van een CQI meetinstrument wordt bepaald met behulp van Principale Componenten Analyse met een oblique rotatie (vanwege de veronderstelde samenhang tussen de factoren). De betrouwbaarheid van de schalen wordt uitgedrukt in de Cronbach's alpha (α), een maat die de interne consistentie van de schaal weergeeft. Een schaal is voldoende betrouwbaar als de $\alpha > 0,70$. Bij een betrouwbaarheid tussen de 0,60 en de 0,70 wordt de schaal voorlopig geaccepteerd. In figuur 2 staan de schalen van de in deze lijst besproken CQI vragenlijsten weergegeven. De Cronbach's alpha (α) van deze schalen varieerde van 0,65 tot 0,94.

Bij de analyses wordt ook gekeken naar het belang dat patiënten/cliënten hechten aan de verschillende kwaliteitsaspecten. Niet alleen kan op basis van de belangscore mede bepaald worden of een item wel of niet uit de lijst verwijderd moet worden, maar ook wordt door een combinatie van belangscores met ervaringscores nagegaan op welke punten de zorg verder verbeterd kan worden of juist al heel goed verloopt. Uit een vergelijking van de belangscores van vijf aandoeningsspecifieke CQI vragenlijsten¹⁷ weten we dat de prioriteiten van patiënten per aandoening variëren. Zo vinden patiënten met een borstafwijking de toegankelijkheid van de zorg en de snelheid waarmee ze terecht kunnen voor onderzoek en behandeling heel belangrijk, terwijl voor patiënten met een chronische aandoening de informatie (over medicatie en behandeling) meer prioriteit heeft. Ook patiënten met een heup- of knie-operatie hechten veel belang aan goede informatievoorziening. Deze verschillen worden weerspiegeld in het algemene waarderingscijfer.

In de rapporten van de ontwikkeltrajecten worden op basis van de belangscores en de gerapporteerde ervaringen van patiënten / cliënten verbeterscores berekend. In tabel 1 wordt een voorbeeld gegeven van de vijf kwaliteitsaspecten die het meest voor verbetering in aanmerking komen bij de zorg rondom Reumatoïde Artritis. De verbeterscores kunnen variëren tussen 1 en 4. Een hogere verbeterscore betekent dat er vanuit patiëntenperspectief de meeste kwaliteitswinst is te behalen.

[FIGUUR 2]

[TABEL 1]

TEST VAN HET DISCRIMINEREND VERMOGEN

Een belangrijk doel van de CQ-index is het zichtbaar maken van verschillen in kwaliteit van zorg tussen verschillende instellingen, zodat patiënten daar hun keuzes op kunnen

baseren en zorgverzekeraars er bij hun zorginkoop op kunnen sturen. Daarom worden de instrumenten getest op hun discriminerend vermogen. Belangrijke voorwaarde daarbij is een voldoende grote steekproef: als vuistregel geldt 20 instellingen of praktijken met elk 200 patiënten/cliënten. De test zelf wordt uitgevoerd door middel van multilevel regressieanalyses, waarbij zonodig gecontroleerd wordt voor variabelen in de steekproef die verschillen tussen de instellingen kunnen veroorzaken (zogenaamde 'case-mix variabelen'). In het artikel van Van Plaggenhoef et al.¹⁸ elders in dit themanummer wordt nader ingegaan op de techniek van de multilevel analyses.

Bij de test op het discriminerend vermogen van de CQI Heup/knie-operatie kwamen weinig verschillen tussen ziekenhuizen naar voor. Dit gold zowel voor de schaalesscores als voor de beoordelingscijfers en de afzonderlijke items die onderzocht werden. Een gering discriminerend vermogen kan duiden op een geringe sensitiviteit van het instrument om verschillen vast te stellen, maar ook op het feit dat er mogelijk niet zoveel verschillen zijn in kwaliteit van zorg vanuit het perspectief van de patiënt die een heup- of knie-operatie ondergaat. Dit laatste lijkt waarschijnlijk. Wanneer we kijken naar de CQI Huisartsenzorg zijn daar wél verschillen tussen de praktijken waarneembaar (tabel 2). Een gebruikelijke manier om verschillen weer te geven is in de vorm van een sterrentabel, waarbij twee sterren een gemiddelde kwaliteit van zorg betekenen, één ster onder- en drie sterren bovengemiddeld.

Behalve om na te gaan of het instrument in staat is verschillen tussen aanbieders vast te stellen worden in deze fase soms ook nog wijzigingen in de formulering van vragen of in de opbouw van de lijst uitgetest. Uiteindelijk resulteert deze fase in een definitief voorstel voor een CQI meetinstrument, dat met de goedkeuring van de werkgroep aan de Wetenschappelijke Adviesraad van het Centrum Klantervaring Zorg kan worden voorgelegd.

[Tabel 2.]

CONCLUSIES EN DISCUSSIE

Bij de ontwikkeling van een CQ-index wordt een vaste procedure gevolgd, waarbij de inbreng van patiënten/cliënten en andere 'stakeholders' gewaarborgd is. De gestandaardiseerde meetsystematiek maakt onderlinge vergelijking tussen aanbieders, sectoren en doelgroepen mogelijk. De CQ-index vragenlijsten verschillen in de mate waarin zij discrimineren tussen zorginstellingen en -aanbieders. De vraag is of bij een gering discriminerend vermogen de vragenlijst te weinig sensitief is of dat er in de zorg niet zulke grote verschillen zijn. Om dit vast te kunnen stellen is er behoefte aan onderzoek waarmee de externe validiteit van de CQI wordt bepaald. Aanvankelijk werd de mogelijkheid om aparte lijsten per sector, groep aanbieders, aandoening of medische ingreep gezien als een van de voordelen van de CQ-index. Inmiddels rijst de vraag hoeveel lijsten er eigenlijk ontwikkeld moeten worden. Momenteel zijn er zeven lijsten ontwikkeld en zijn er nog bijna 20 in voorbereiding. Voor een deel overlappen de lijsten elkaar qua inhoud: sectoren en aanbieders vormen een matrix met aandoeningen en medische ingrepen. Niet alleen qua thematiek en schalen, maar ook de doelgroep is gedeeltelijk hetzelfde. Datzelfde geldt voor sectorgerichte lijsten en lijsten die zich richten op een ketentraject. Het Centrum Klantervaring Zorg speelt in deze een belangrijke rol aangezien zij het totale veld aan ontwikkeltrajecten overziet. Bovendien adviseert het Centrum omtrent de financiering van ontwikkeltrajecten uit publieke middelen. Na afronding van alle genoemde ontwikkeltrajecten is voor de belangrijkste sectoren een CQ-index beschikbaar. Een aantal aandoeningen met een grote ziektelast (hart- en vaatziekten, diverse vormen van kanker) staan nog op de prioriteitenlijst van het Centrum Klantervaring Zorg. Voor andere ziekten of ingrepen zal het moeilijk zijn om financiering te vinden, terwijl de betekenis van goede zorg en keuzemogelijkheden voor deze patiënten net zo wezenlijk is en transparantie in de zorg ook voor hen een recht is.

Standaardisatie van vragen en antwoordcategorieën is belangrijk om optimale vergelijking tussen zorgaanbieders mogelijk te maken, maar niet alle thema's passen in het format. Inmiddels gaan experimenten van start waarin incidenteel andere antwoordcategorieën worden toegestaan. Daarbij kan het gaan om 3-punts schalen (ja, helemaal; ja, gedeeltelijk; nee) of de antwoordcategorieën van de oorspronkelijke QUOTE-vragenlijsten (ja; eigenlijk wel; eigenlijk niet; nee). Belangrijk blijft dat uiteindelijk de resultaten van deze experimenten zullen leiden tot nieuwe richtlijnen die voor alle CQI meetinstrumenten van toepassing zijn.

Hoewel er geen problemen met de respons zijn is de lengte van de vragenlijsten toch een thema dat bij herhaling terugkeert. Het zou goedkoper zijn om kortere lijsten te kunnen uitzetten, zeker ook in reguliere metingen die na de ontwikkeling uitgevoerd gaan worden door verzekeraars en aanbieders. Bovendien worden respondenten door kortere lijsten minder belast. De oorzaak van de lengte ligt voor een deel in het feit dat over een lijst consensus met alle 'stakeholders' moet worden bereikt, omdat de lijst voor meerdere doelen geschikt moet zijn. Daarom is inmiddels een project gestart over de ontwikkeling van doelspecifieke lijsten (zie Triemstra et al.¹⁹ elders in dit themanummer). Deze kortere lijsten zouden gericht zijn op het genereren van data voor keuze-informatie en zorginkoop of op kwaliteitsverbetering. Daarmee zou de CQI systematiek echter wel afstappen van het adagium 'Eén keer meten' en lopen respondenten de kans om lijsten voor verschillende doelen te moeten invullen.

Een risico van een gestandaardiseerde manier van meten waarbij ook vergelijkingen tussen lijsten mogelijk zijn is schijnobjectiviteit. Met name waarderingcijfers op een schaal van 0-10 worden te gemakkelijk als 'objectieve' maat van kwaliteit overgenomen, terwijl uit een vergelijking bij vijf aandoeningsspecifieke lijsten bleek dat deze cijfers uit andere kwaliteitsaspecten waren opgebouwd, omdat de verschillende doelgroepen andere zaken belangrijk vonden bij de beoordeling van de zorg.

Een principieel punt rondom de CQ-index blijft: is een patiënt in staat om kwaliteit van zorg te beoordelen? Hierbij moet worden opgemerkt dat de gegevens uit de CQ-index nooit de enige informatiebron over kwaliteit kunnen zijn: er zijn immers andere indicatoren (van professionele aard of ten aanzien van de bedrijfsvoering) waar een patiënt geen zicht op heeft. Op dit moment ligt de nadruk in CQI meetinstrumenten vooral op structuur- en procesaspecten in de zorg (toegankelijkheid, bereikbaarheid, bejegening, communicatie, informatie), die belangrijke onderdelen zijn van kwaliteit van zorg.²⁰ Wel wordt vooral vanuit verzekeraars de vraag naar zogenaamde Patient Reported Outcome Measures (PROM's) in CQI meetinstrumenten groter. Op dit moment wordt onderzocht hoe deze het meest valide en betrouwbaar gemeten kunnen worden. De legitimering van de CQ-index ligt echter toch vooral in de kwaliteitsaspecten die voor patiënten en cliënten zelf het belangrijkste zijn bij de beoordeling van en het maken van keuzes in de zorg.

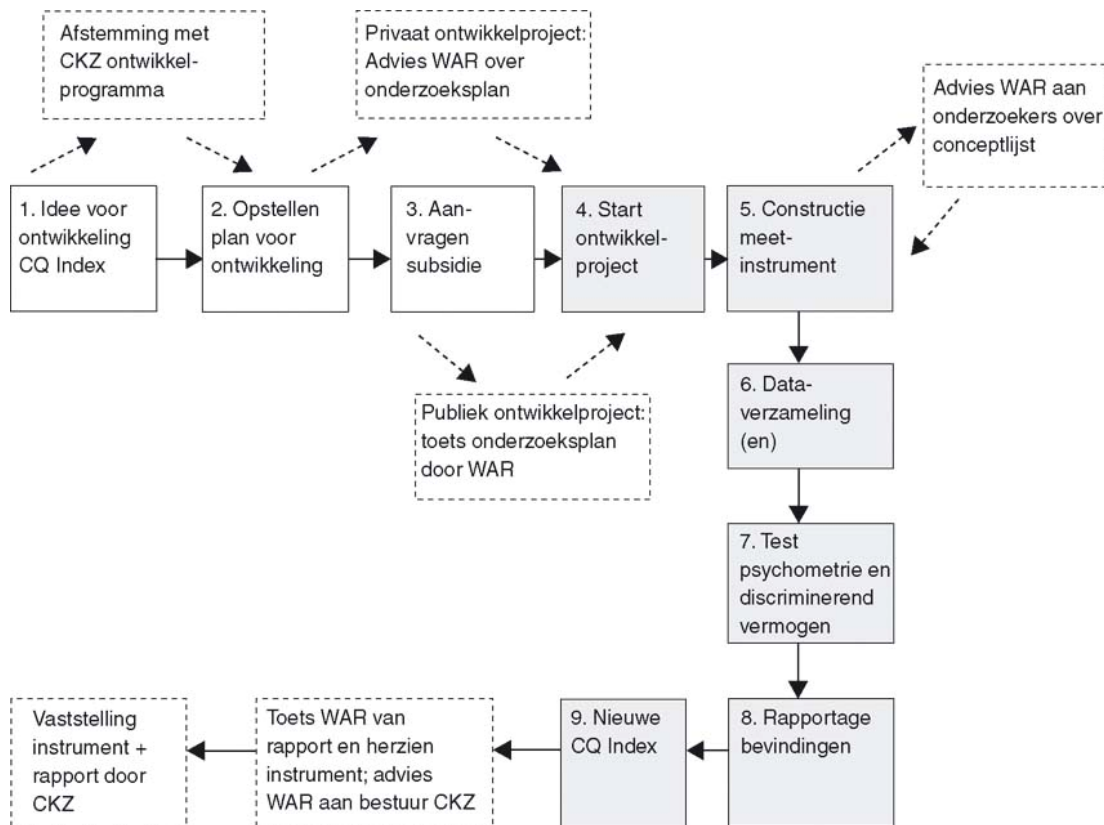
LITERATUUR

1. Campen C van, Sixma HJ, Kerssens JJ, Peters L. Assessing noninstitutionalized asthma and COPD patients' priorities and perceptions of quality of health care: the development of the QUOTE-CNSLD instrument. *J Asthma* 1997;34:531-8.
2. Campen C van, Sixma HJ, Kerssens JJ, Peters L, Rasker JJ. Assessing patients' priorities and perceptions of the quality of health care: the development of the QUOTE-Rheumatic-Patients instrument. *Br J Rheumatol* 1998;37:362-8.
3. Sixma HJ, Kerssens JJ, Campen C van, Peters L. Quality of care from the patients' perspective: from theoretical concept to a new measuring instrument. *Health Expect* 1998;1:82-95.
4. Sixma HJ, Campen C van, Kerssens JJ, Peters L. Quality of care from the perspective of elderly people: the QUOTE-elderly instrument. *Age Ageing* 2000;29:173-8.
5. Cleary PD, Edgman-Levitan S. Health care quality. incorporating consumer perspectives.

- JAMA. 1997;278:1608-12.
6. Sixma HJ, Delnoij DMJ. (red). Handboek CQI-meetinstrumenten; een handleiding voor de ontwikkeling en het gebruik van Consumer Quality Index (CQI) meetinstrumenten. Utrecht: b, 2007 (www.centrumklantervaringzorg.nl; 4 juni 2008).
 7. Meuwissen L, Bakker D de. CQ-index huisartsenzorg: meetinstrumentontwikkeling. De ontwikkeling van het instrument, de psychometrische eigenschappen en het discriminerend vermogen. Utrecht: NIVEL, 2008.
 8. Gelsema T, Delnoij DMJ, Triemstra, AHM. Kwaliteit van zorg rondom een totale heup- of knieoperatie vanuit het perspectief van patiënten: meetinstrumentontwikkeling en validatie. Utrecht: NIVEL, 2006.
 9. Stubbe J, Dijk L van. Het discriminerend vermogen van de CQ index Heup-/Knieoperatie. Utrecht: NIVEL, 2007.
 10. Damman OC, Hendriks M, Triemstra AHM, Sixma HJ. CQ-index Mammacare: meetinstrumentontwikkeling. Kwaliteit van zorg rondom een borstafwijking vanuit het perspectief van patiënten. Utrecht: NIVEL, 2007.
 11. Koopman L, Rademakers JJDJM. Discriminerend vermogen van de CQ-index Mammacare. Utrecht: NIVEL, 2008.
 12. Zuidgeest M, Hendriks M, Triemstra AHM, Sixma HJ. CQ-index Reumatoïde Artritis: meetinstrumentontwikkeling. Kwaliteit van reumazorg vanuit het perspectief van patiënten met reumatoïde artritis. Utrecht: NIVEL, 2007.
 13. Koopman L, Rademakers JJDJM. Discriminerend vermogen van de CQ-index Reumatoïde Artritis. Utrecht: NIVEL, 2008.
 14. Zuidgeest M, Boer D de, Hendriks M, Rademakers J. Verschillende dataverzamelmethode in CQI onderzoek: een overzicht van de respons en representativiteit van respondenten. Tijdschr Gezondheidswet 2008;86:455-62.
 15. Morgan DL. The focus group kit (1): The focus group guide book. Thousand Oaks : Sage, 1998
 16. Triemstra AHM, Rademakers JJDJM. Focus groups as a tool for developing measurement instruments from a patient's perspective. European Journal of Public Health 2008; 18, Supplement 1:58.
 17. Hendriks M, Boer D de, Spreeuwenberg P, Rademakers J, Delnoij D. De CQ-index: vergelijken van klantervaringen in de zorg. Utrecht: NIVEL/CKZ, 2008.
 18. Plaggenhoef W van, Vunderink L, Weert C van, Vriens B. Het effect van ziekenhuiskenmerken op het rapportcijfer voor de zorg. Tijdschr Gezondheidswet 2008;86:478-85.
 19. Triemstra AHM, Hendriks M, Delnoij DMJ, Rademakers JJDJM. Doelspecifieke versies van CQ-index meetinstrumenten: korter, krachtiger, en specifiekere meten? Tijdschr Gezondheidswet 2008;86:486-93.
 20. Donabedian A. Explorations in quality assessment and monitoring (vol 1): the definition of quality and approaches to its assessment. Michigan, Ann Arbor: Health Administration Press, 1980.

FIGUREN EN TABELEN

Figuur 1. Proces van ontwikkeling en vaststelling van een CQ-index vragenlijst



Figuur 2. Schalen die na de psychometrische tests werden gevormd in de CQI meetinstrumenten Bron: ^{7,8,10,12}

CQI Huisartsenzorg	CQI Heup/knie-operatie
toegankelijkheid praktijkassistente	bejegening verpleegkundigen
bejegening huisarts	hulp van verpleegkundigen
communicatie huisarts	bejegening artsen
zorg op maat huisarts	Informatie
samenwerking andere zorgverlener	pijnbestrijding
bejegening andere zorgverlener	medicatie
zorg op maat andere zorgverlener	huisarts
CQI Mammacare	CQI Reumatoïde Artritis
bejegening zorgverleners tijdens onderzoeken	bejegening reumatoloog tijdens eerste bezoek
bejegening huisarts	bejegening zorgverlener tijdens controle van zorg
bejegening verpleegkundigen fase operatie	deskundigheid zorgverlener tijdens controle van zorg
bejegening chirurg fase operatie	bejegening medicijnverstrekker
autonomie / uitleg / vragen stellen	bejegening gespecialiseerd verpleegkundige
autonomie / meebeslissen behandelinge	deskundigheid gespecialiseerd verpleegkundige
bejegening zorgverleners radiotherapie	bejegening therapeut
informatievoorziening radiotherapie	bejegening psychosociale hulpverlener
bejegening zorgverleners chemotherapie	bejegening bedrijfsarts
informatievoorziening chemotherapie	afstemming
afstemming en samenwerking	
toegankelijkheid van de zorg	
psychosociale zorg	
fysiotherapie	
revalidatie	

Tabel 1 Top 5 verbeterscores (V) met bijbehorende belangscores (B) en het percentage negatieve ervaringen (E) (CQI Reumatoïde Artritis)¹²

Nr.	Kwaliteitsaspect	B	E	V
110	Informatievoorziening over RA via een speciale voorlichtingswebsite van het ziekenhuis	2,61	82	2,11
122	Informatie over vergoedingen en eigen bijdragen van hulp en hulpmiddelen	3,31	51	1,69
5	Huisarts heeft gevraagd naar stijve gewrichten	2,98	54	1,61
119	Informatie over vergoedingen uit aanvullende verzekeringen van verschillende zorgverzekeraars	3,32	48	1,60
116	Weten waar een fout gemeld kan worden als er iets niet goed is gegaan	3,32	48	1,59

Nummers overeenkomstig met ervaringenlijst

Belangscores (B) worden gemeten met een 4-puntsschaal: 1= niet belangrijk; 4 = van het aller grootste belang

Bij E wordt het percentage respondenten vermeld dat aangeeft deze informatie of zorg nooit of soms gekregen te hebben

Tabel 2 Aantal praktijken (N=32) met één, twee en drie sterren voor ervaringen met de zorg uit de huisartsenpraktijk⁷

Ervaringsscores en waarderingscijfer*	1 ster	2 sterren	3 sterren
Schaal toegankelijkheid	13	5	14
Schaal assistente	10	9	13
Schaal 'zorg op maat' huisarts	9	13	10
Schaal communicatie en bejegening huisarts	9	14	9
a. schaal bejegening huisarts	9	14	9
b. schaal communicatie huisarts	8	15	9
Schaal samenwerking en bejegening andere zorgverlener	2	27	3
a. schaal samenwerking andere zorgverlener	2	26	4
b. schaal bejegening andere zorgverlener	3	25	4
Schaal 'zorg op maat' andere zorgverlener	4	25	3
Algemeen: het cijfer van 0 tot 10 dat aan de huisartspraktijk wordt gegeven	10	12	10

* Hierbij is gecorrigeerd voor opleiding, geslacht, gezondheid en leeftijd

Toelichting: praktijk met score 3 sterren hebben significant beter gescoord dan praktijken met 1 ster.