

Ziekenhuizentop-100: wisselende ranglijsten, wisselende reputaties

R.W.M.Giard

'Het grootste gedeelte van de hulpverlening bestaat niet uit transacties waarbij een dokter iets aan een patiënt verkoopt, maar uit interacties.'¹

Maakt het voor de kansen op herstel uit of een patiënt met een bepaalde kwaal ziekenhuis A dan wel ziekenhuis B binnenstapt? En waar vindt die patiënt de informatie om tussen zorginstellingen te kunnen kiezen? Diverse instanties zetten zich ervoor in de kwaliteit van zorg transparant te maken door kwaliteitsindicatoren te publiceren. Zo verschenen er in de maand september in de publieksmedia maar liefst drie verschillende ranglijsten van ziekenhuizen: die van het *Algemeen Dagblad* (AD),² van het bureau Roland Berger Strategy Consultants, gepresenteerd in het televisieprogramma NOVA,³ en van *Elsevier*.⁴ De minister van Volksgezondheid juicht het transparant maken van de zorg toe; hij ziet deze initiatieven als een vooruitgang. Het past immers goed in het nieuwe klimaat van marktwerking in de zorg.

Overzichten over uitkomsten van medische zorg worden al sinds de 19e eeuw gepubliceerd. Op aandringen van Florence Nightingale werden gegevens verzameld om de mortaliteit te kunnen beoordelen in relatie tot de aandoening en de daarvoor gegeven behandeling, zodat eventueel verbeteringen in de zorg konden worden aangebracht. In 1862 publiceerde de Royal Statistical Society of London voor het eerst dergelijke uitkomsten.⁵

Naar het voorbeeld van het onderzoek van scholen door het dagblad *Trouw* zijn diverse media begonnen met het jaarlijks publiceren van vergelijkend kwaliteitsonderzoek van Nederlandse ziekenhuizen. Wij zien hier dus intern onderzoek door het eigen ziekenhuis tegenover door buitenstaanders uitgevoerd breed kwaliteitsonderzoek.

Geven dergelijke onderzoeken een betrouwbaar beeld van de kwaliteit van de ziekenhuizen? Worden zorginstellingen met een minder goede score geprikkeld om zinvolle en vruchtbare verbeteracties te ondernemen? In hoeverre laten patiënten zich leiden door deze onderzoeken? Aan de hand van deze vragen zal ik deze ontwikkeling binnen de Nederlandse gezondheidszorg tegen het licht houden.⁶

MOTIEVEN EN METHODEN VOOR HET METEN VAN DE KWALITEIT VAN ZORG

Een landelijk vergelijkend onderzoek van de kwaliteit van ziekenhuizen is gemakkelijker gezegd dan gedaan.^{7 8} Waarom zou een dergelijke grootschalige externe kwaliteitstoetsing nodig zijn, hoe moet die worden uitgevoerd en op welke wijze worden de uitkomsten gepresenteerd?

Goede kwaliteit is nooit vanzelfsprekend en daarom is voortdurende controle nodig. Een van de journalistieke taken is het dienen van een publiek belang door het signaleren van ongewenste situaties. Een landelijk onderzoek naar mogelijke verschillen in ziekenhuiskwaliteit legt deficiënties bloot. Voorts wordt het patiënten mogelijk gemaakt aan de hand van de resultaten bewust tussen ziekenhuizen te kiezen.⁹

Vervolgens kom ik op de uitvoering van een dergelijk onderzoek. Kwaliteit van zorg dient eerst gedefinieerd en vervolgens gemeten te worden.^{7 9 10} Kwaliteit zou voor dit type onderzoek kunnen worden omschreven als de mate waarin zorginstellingen in staat blijken de gewenste en volgens de stand van de medische wetenschap te bereiken uitkomsten bij patiënten te realiseren.^{7 9}

Aangezien de media vanuit het perspectief van patiënten werken, worden tenminste twee dimensies onderzocht: de klanttevredenheid, hetgeen een combinatie van bejegening en patiëntenlogistiek is, en de medische prestaties. Wordt er bij de medische prestaties dan vooral gekeken naar hoe artsen handelen (procestoetsing), naar de resultaten daarvan (uitkomsttoetsing) of naar allebei?^{10 11} Bij extern onderzoek moet men kunnen beschikken over de juiste gegevens die een betrouwbaar oordeel over kwaliteit mogelijk maken.

Als wij ziekenhuizen qua prestaties met elkaar willen vergelijken, zijn ze dan ook werkelijk vergelijkbaar? Wanneer verschillen worden geconstateerd, is dat dan op basis van toeval, op basis van bias, indien door selectiemechanismen uiteenlopende soorten patiëntenpopulaties werden vergeleken, of gaat het toch om verschillen in prestatie? De constatering dat de kwaliteit van zorg in een ziekenhuis onder het bereikbare optimum ligt, is een diagnose per exclusionem: eerst dienen alle andere mogelijke verklaringen te zijn uitgesloten. Bij genoemde vergelijkende onderzoeken kreeg elke zorginstelling een puntenaantal. Het is echter de vraag of de kwaliteit van een ziekenhuis in één getal is te vangen.

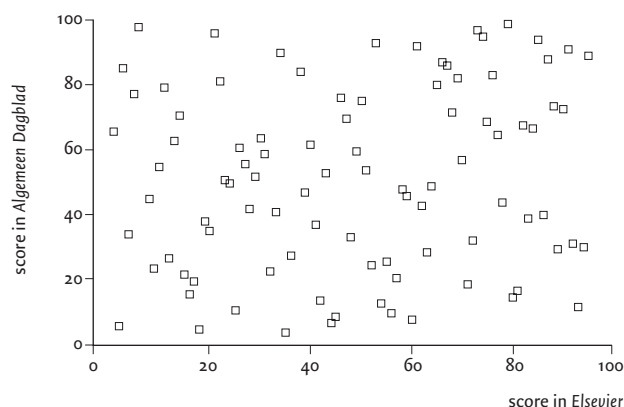
Medisch Centrum Rijnmond-Zuid, locatie Clara, afd. Klinische Pathologie, Postbus 9119, 3007 AC Rotterdam.
Hr.dr.R.W.M.Giard, patholoog, klinisch epidemioloog en jurist (tevens: Izkazia Ziekenhuis, afd. Pathologie, Rotterdam) (giardr@mcrz.nl).

De drie studies rangschikken de Nederlandse ziekenhuizen in volgorde van kwaliteit, waarbij winnaars en verliezers worden aangegeven. Alle drie hebben ze duidelijk één ding gemeen: men pretendeert onderscheid te kunnen maken tussen kwalitatief goede en minder goede zorginstellingen. Qua methode is bij alle drie gekozen voor een multifactoriële analyse met puntentellingen en een uiteindelijke optelscore als basis voor de rangschikking. Qua praktische uitvoering verschillen de onderzoeken onderling echter sterk. Het AD baseert de uitkomsten op 30 uiteenlopende medische en organisatorische toetsingscriteria (www.ad.nl/ziekenhuistop100). Elsevier gaat uit van antwoorden van geselecteerde informanten op vragenlijsten en men onderzoekt drie rubrieken: medische en verpleegkundige prestaties, specialistische zorg en tenslotte organisatie en service. Elke rubriek bevat verschillende onderwerpen met in totaal 54 scoreonderdelen. Roland Berger Strategy Consultants gebruikt door de Inspectie voor de Gezondheidszorg gepubliceerde prestatie-indicatoren, aangevuld met gegevens uit de jaarverslagen van ziekenhuizen.

AD en Elsevier publiceren een ranglijst van alle ziekenhuizen, Roland Berger Strategy Consultants presenteert een rangschikking van alle ziekenhuizen, exclusief de academische ziekenhuizen en een oncologisch centrum, in vier kwartielen van de 25% beste, de 25% slechtste en twee tussenliggende categorieën. Elsevier geeft daarnaast op grond van de totaalscore een groepering van de ziekenhuizen in de categorieën uitstekend-goed-gemiddeld-matig-slecht.

Bieden de drie verschillende ranglijsten de patiënt ondanks hun verschillende benaderingen een eenduidig beeld van de kwaliteit van de onderzochte ziekenhuizen? In figuur 1 is de correlatie tussen de scorelijsten van AD en Elsevier weergegeven. Uit de spreiding van de punten wordt duidelijk dat er vrijwel geen correlatie tussen beide beoordelingen bestaat. Soms zijn de tegenstellingen zelfs groot: een ziekenhuis kan bij het ene onderzoek tot de categorie beste ziekenhuizen behoren en bij het andere juist tot de slechtste.

Als in het onderzoek van Elsevier de classificaties 'uitstekend' en 'goed' worden samengevoegd en de vier rubrieken vervolgens worden vergeleken met de vier kwartielen van Roland Berger Strategy Consultants, blijkt eveneens dat er nauwelijks overeenstemming tussen beide bestaat (tabel). Wanneer alle drie onderzoeken onderling op het punt van waarnemerovereenstemming worden onderzocht, zijn er drie combinatieparen (figuur 2). Uitgedrukt in kappacoëfficiënten is de score hoogstens 0,14 en dat betekent in alle gevallen slechte overeenstemming ten aanzien van de oordelen.¹²



FIGUUR 1. Correlatie tussen de ziekenhuisranglijsten van het Algemeen Dagblad en Elsevier (correlatiecoëfficiënt $r = 0,14$).^{2, 4} Bij volledige overeenstemming van de beoordelingen zouden de punten op een rechte, min of meer stijgende lijn liggen ($r = +1$).

DE VOORS EN TEGENS VAN HET PUBLICEREN VAN RANGLIJSTEN

Het publiek maken van uitkomsten van kwaliteitstoetsing, of dit nu gebeurt door toezichthouders of de media, is geen doel op zich. Het gaat immers om goede patiëntenzorg. De vraag dringt zich op of publiek maken een effectieve manier is om de zorg te verbeteren. Ranglijsten zouden heilzaam kunnen werken doordat ze slecht scorende ziekenhuizen prikkelen om tot betere prestaties te komen. Er is weinig experimenteel onderzoek uitgevoerd om hiervoor overtuigende argumenten te leveren.^{9, 13} Maar hier gaat nog een kwestie aan vooraf: op welke objectieve en reproduceerbare wijze kan vergelijkenderwijs de kwaliteit van ziekenhuizen in kaart gebracht worden? Daarbij moet aan verschillende

Variatie in beoordelingen van 85 ziekenhuizen door Elsevier en Roland Berger Strategy Consultants;^{3, 4*} weergegeven zijn aantallen ziekenhuizen

beoordeling in Elsevier	beoordeling door Roland Berger Strategy Consultants in kwartielen				totaal
	1 (beste)	2	3	4 (slechtste)	
goed of uitstekend	7	8	11	3	29
gemiddeld	11	7	4	9	31
matig	2	3	4	7	16
slecht	1	3	2	3	9
totaal	21	21	21	22	85

* κ voor overeenstemming tussen beide beoordelingen: 0,07 (95%-BI: -0,005--0,21) (zie ook figuur 1). Waarden voor κ variëren van 0 tot 1,00: bij 0 bestaat er alleen toevallige overeenstemming en bij 1,00 bestaat er volledige overeenstemming.

voorwaarden worden voldaan: de uitkomsten dienen overtuigend én – bij meerdere onafhankelijke onderzoeken – eenduidig te zijn. Dan pas krijgt men een geloofwaardig pedagogisch signaal.¹⁴

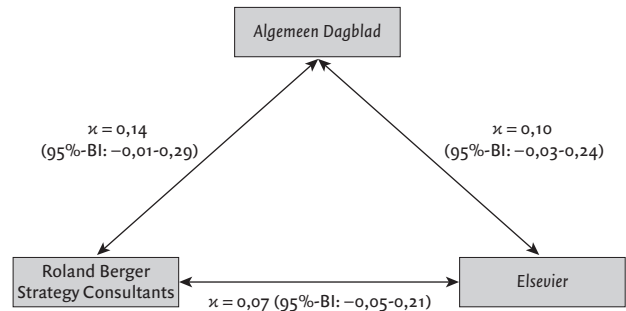
Het openbaar maken van de kwaliteit van een ziekenhuis vindt zijn grondslag in het recht op informatie binnen een democratische samenleving. De geïnformeerde burger kan kiezen. Maar wat doen patiënten met de vergelijkende studies? Vooralsnog weinig, zo blijkt uit onderzoek.^{9 13 14} In acute situaties of vanwege zorgpolissen zijn de keuzemogelijkheden voor de patiënt praktisch veelal beperkt.

Naast positieve effecten kan publicatie van kwaliteitsbeoordelingen ook averechtse effecten opleveren. Alleen al het definiëren van 'voortreffelijkheid', het meten daarvan en het verspreiden van de resultaten beïnvloeden het systeem: men richt de aandacht op de gemeten variabelen.¹⁵ Diverse auteurs wijzen op de volgende disfunctionele consequenties van het publiceren van gegevens over kwaliteit: de ontwikkeling van strategisch gedrag om de getallen te flatteren; het zich vooral richten op kortetermijndoelen; voorzichtigheid bij het doorvoeren van innovaties; de ontwikkeling van een tunnelvisie: alleen de gemeten onderwerpen veranderen, niet de niet-beoordeelde; het te veel nadruk leggen op de belangen van individuele organisaties, niet van de totale medische zorg; en het 'creatief' registreren van data of zelfs het voorkomen van fraude.^{9 13-15}

Als het gaat om de betekenis van onderzoeken die prestaties van ziekenhuizen vergelijken, wordt één ding duidelijk: er zijn meer verwachtingen dan feiten en er ligt dus een forse researchagenda te wachten.^{9 11 13}

Vanwege de uiteenlopende methoden die zijn toegepast, zijn de drie ranglijsten niet goed te vergelijken, maar de onderzoekers pretenderen wel een kwaliteitsklassement te kunnen maken. Aangezien er over de meeste ziekenhuizen in de verschillende rapporten uiteenlopend wordt geoordeeld, gaat daarvan niet een eenduidig pedagogisch signaal uit. Evenmin krijgt de patiënt een helder beeld welk ziekenhuis hij nu het beste kan kiezen.

Is zo'n totaalscore per ziekenhuis, waarin medisch inhoudelijke, organisatorische, logistieke, economische en communicatieve aspecten zijn samengevoegd, wel bruikbaar? De ongeveer 100 Nederlandse ziekenhuizen zijn in vele opzichten onderling verschillend. Primair gaat het om de juiste aanpak van het medische probleem. Daarbij komt de door Donabedian geformuleerde trias structuren-processen-uitkomsten goed van pas.^{8 16} Steeds meer blijkt uit empirisch onderzoek dat het systematisch onderzoeken van alle elementen van de trias zowel voor meting als voor verbetering van kwaliteit het beste is; uitkomsten alleen vormen geen goede kwaliteitsmaat.^{7 11} Recente onderzoeken, toegespitst op veelvoorkomende medische problemen en geschoeid op Donabedians leest, waarbij zorgkwaliteit



FIGUUR 2. Onderlinge overeenstemming tussen 3 ranglijsten van ziekenhuizen.²⁻⁴ De beoordelingen werden ondergebracht in 4 rubrieken van beste naar slechtste en de aangegeven κ -waarden werden berekend na lineaire weging. Waarden voor κ variëren van 0 tot 1,00: bij 0 bestaat er alleen toevallige en bij 1,00 bestaat er volledige overeenstemming.

tussen ziekenhuizen kon worden vergeleken, laten zien dat deze benadering veel zinvoller is.^{17 18}

Een hoogleraar Gezondheidswetenschappen verklaarde naar aanleiding van de AD-ranglijst dat hij zich liever niet zou laten behandelen in de slechtste 20 ziekenhuizen.² Er zijn met dergelijke onderzoeken dus reputaties van ziekenhuizen in het geding. Als het inderdaad uitmaakt of een patiënt met een bepaalde kwaal ziekenhuis A dan wel ziekenhuis B binnenstapt, kunnen de medische professionals daarin niet berusten: er wordt een appel gedaan op de beroepseer als motivatiebron en niet op koopmanschap.

Financiële ondersteuning: geen gemeld. Belangenconflict: de auteur werkt in twee ziekenhuizen, waarvan in de drie beschreven ranglijsten het ene als 'uitstekend' of 'middelmatig' is geïnclassificeerd en het andere consequent als 'slecht'.

Aanvaard op 9 oktober 2006

Literatuur

- 1 Roovers D. Bij de huisarts: 'Je moet daadkrachtig én gelaten zijn'. Annemarie Mol: het gaat in de spreekkamer niet om autonomie. *Filosofie Magazine*. 2006;nr 5:22-5.
- 2 Ziekenhuizen top-100. *Algemeen Dagblad* 19 september 2006.
- 3 Collot d'Escury J, Alma R, Reenen M van, Remmen T van. *Steering the right course. Dutch hospitals 2006 – key developments and trends*. Amsterdam: Roland Berger Strategy Consultants; 2006.
- 4 Leeuwen A van, Wansink W. *Gezondheidszorg: de beste ziekenhuizen*. Elsevier. 2006;62:83-111.
- 5 Turner T. *Statistics of the general hospitals of London, 1861*. *Journal of the Statistical Society*. 1862;25:384-8.

- 6 Everdingen JJE van. De 'beste' ziekenhuizen van Nederland; kanttekeningen bij de Elsevier-enquête. *Ned Tijdschr Geneeskd.* 2000;144:2537-40.
- 7 Lohr KN. How do we measure quality? *Health Aff (Millwood).* 1997;16:22-5.
- 8 Shojania KG, Showstack J, Wachter RM. Assessing hospital quality: a review for clinicians. *Eff Clin Pract.* 2001;4:82-90.
- 9 Mason A, Street A. Publishing outcome data: is it an effective approach? *J Eval Clin Pract.* 2006;12:37-48.
- 10 Giard RWM. Prestatie-indicatoren als maat voor de kwaliteit van medische zorg: retoriek en realiteit. *Ned Tijdschr Geneeskd.* 2005;149:2715-9.
- 11 Jha AK. Measuring hospital quality: what physicians do? How patients fare? Or both? *JAMA.* 2006;296:95-7.
- 12 Fleiss JL. *Statistical methods for rates and proportions.* New York: Wiley; 1981. p. 218.
- 13 Marshall MN, Romano PS. Impact of reporting hospital performance. *Qual Saf Health Care.* 2005;14:77-8.
- 14 Werner RM, Asch DA. The unintended consequences of publicly reporting quality information. *JAMA.* 2005;293:1239-44.
- 15 Wachter RM. Expected and unanticipated consequences of the quality and information technology revolutions. *JAMA.* 2006;295:2780-3.
- 16 Donabedian A. The quality of care. How can it be assessed? *JAMA.* 1988;260:1743-8.
- 17 Bradley EH, Herrin J, Elbel B, McNamara RL, Magid DJ, Nallamothu BK, et al. Hospital quality for acute myocardial infarction: correlation among process measures and relationship with short-term mortality. *JAMA.* 2006;296:72-8.
- 18 Mohammed MA, Mant J, Bentham L, Raftery J. Comparing processes of stroke care in high- and low-mortality hospitals in the West Midlands, UK. *Int J Qual Health Care.* 2005;17:31-6.

Abstract

Top 100 hospitals: varying lists, varying reputations. – In the month of September no fewer than 3 different hospital rankings appeared in the media. These rankings do not reliably portray the quality of hospitals. Since they mutually differ substantially, they provide insufficient information to motivate hospitals with lower scores to make meaningful improvements, and they only partly influence the choice of patients. Another unfavourable effect of these rankings is that hospitals may begin to focus on measured factors at the expense of unmeasured factors, such as innovation, long-term goals and cooperation. Improvements in patient care can be expected to come only from the ethics of medical professionals. *Ned Tijdschr Geneeskd.* 2006;150:2355-8