

A photograph of a row of empty, blue, airport-style chairs with silver metal frames and black armrests. The chairs are arranged in a line, receding into the background. The floor is a light-colored, polished material that reflects the chairs. The background is slightly blurred, showing a white wall and a blue door or partition.

**HET BESTE
ZIEKENHUIS
BESTAAT NIET!**

HET BESTE ZIEKENHUIS BESTAAT NIET!

Over prestatie­meting en rankings bij ziekenhuizen

Eind mei 2011 werd bekend dat in het Maasstad Ziekenhuis te Rotterdam op de intensive care een uitbraak was geconstateerd van de klebsiella bacterie. Het betreft hier een bacterie die niet meer gevoelig is voor zelfs de meest zware antibiotica. Voor reeds ernstig zieke patiënten kan dit tot hun overlijden leiden. Na enkele maanden is het gelukt de uitbraak onder controle te krijgen. Na juli 2011 zijn er geen nieuwe gevallen meer geconstateerd. Naast de ernst van de situatie werd bovendien al vrij snel bekend dat het ziekenhuis (of beter: medewerkers van het ziekenhuis) al vele maanden op de hoogte waren van de uitbraak en er door de leiding van het ziekenhuis niet of nauwelijks op is gereageerd.

De uitbraak werd groot nieuws en haalde kranten, radio en televisie. Een en ander heeft uiteindelijk geleid tot het aftreden van de Raad van Bestuur en de Raad van Toezicht van het ziekenhuis. De nieuwe leiding (ad interim) van het ziekenhuis heeft besloten tot het instellen van een zwaar opgetuigd extern onderzoek, onder leiding van oud-SG Lemstra. Eind maart 2012 heeft de commissie Lemstra haar rapport, met als titel 'Oog voor het onzichtbare' opgeleverd. Eén van de meest opvallende constatering­en uit het rapport betreft de rol die de scorelijstjes van het Algemeen Dagblad en Elsevier hebben gespeeld bij de strategie van de leiding van het ziekenhuis. In deze lijstjes wordt aangegeven welke ziekenhuizen in Nederland het beste en het slechtste scoren. Daarbij is tevens opmerkelijk dat het Maasstad Ziekenhuis in 2010 volgens de ranking van het Algemeen Dagblad behoorde tot de vijf beste ziekenhuizen van ons land (plaats nummer 5 in de lijst van de ongeveer honderd Nederlandse ziekenhuizen). Overigens was het ziekenhuis (mede door de uitbraak) gezakt naar nummer 30 en kregen zij in het blad Elsevier twee van de vier punten in 2011, hetgeen aangeeft dat het ziekenhuis middelmatig is.

In deze bijdrage zullen wij ingaan op de betrouwbaarheid van de lijstjes van 'beste' ziekenhuizen en op perverse effecten die daarvan uit kunnen gaan. De centrale vragen in deze bijdrage zijn dan ook of vastgesteld kan worden welk ziekenhuis het beste is en hoe betrouwbaar de rankings (lijstjes) van met name Elsevier en het Algemeen Dagblad zijn. Alvorens dit te behandelen zal eerst nog nader in worden gegaan op het rapport van de commissie Lemstra en op de methodiek van ons onderzoek.

Samenvatting

Enige tijd geleden presenteerde de commissie Lemstra haar rapport over het Maasstad ziekenhuis Rotterdam, waar de klebsiella bacterie heerste. Opvallend uit het rapport is de dominante rol die de lijstjes met de 'beste ziekenhuizen' van Elsevier en het Algemeen Dagblad vervullen binnen de strategie van dit ziekenhuis (en waarschijnlijk bij vele andere). In dit artikel wordt aandacht besteed aan de verschillen, de overeenkomsten en de conclusies die aan deze rankings worden verbonden. De performance bubble wordt doorgeprikt: het beste ziekenhuis bestaat niet. Ten eerste wordt duidelijk dat er geen eenduidige definitie van kwaliteit wordt gehanteerd, de onderzoeken richten zich op verschillende kenmerken van kwaliteit. Daarnaast worden gegevens aangeleverd door de (bestuurders van) ziekenhuizen zelf, wat strategisch gedrag opwekt en de betrouwbaarheid van de gegevens in negatieve zin beïnvloedt. Ten derde blijken er fouten te zitten in de toegekende scores. Kortom, de gegevens zijn deels onbetrouwbaar, de meetmethodes zijn aanvechtbaar en de eindoordelen zijn arbitrair.

Auteurs

Gerrit Dijkstra,
De Haagse Hogeschool &
Universiteit Leiden
Maarten Haverkamp,
De Haagse Hogeschool.

Correspondentieadres

g.s.a.dijkstra@hhs.nl

Lectoraat

Public Management:
verantwoordelijkheidsrelaties
binnen de publieke sector

Lector

Gerrit Dijkstra

Het rapport van de commissie Lemstra

Uit het rapport van de commissie Lemstra wordt duidelijk dat er vele ‘schuldigen’ aan te wijzen zijn, of anders geformuleerd, er is niet een enkele duidelijke schuldige aan te wijzen. In het rapport wordt een aantal verantwoordelijken genoemd, waaronder de medisch microbiologen. De commissie Lemstra constateert dat op hen een zware verantwoordelijkheid rust voor het voortduren en verergeren van de uitbraak. Maar ook de afdeling infectiepreventie is tekort geschoten, hetzelfde geldt voor de betrokken medisch specialisten en de medische staf als geheel. Naast de medici wijst de commissie erop dat ook op de Raad van Bestuur, de directie en de Raad van Toezicht een zware verantwoordelijkheid rust voor de omvang van de uitbraak. In termen van Bovens (1990) is er dan ook sprake van het probleem van de vele handen. Waar velen verantwoordelijk zijn, is uiteindelijk niemand verantwoordelijk.

Over de rol van de Raad van Bestuur is de commissie duidelijk. De aandacht voor kwaliteit en veiligheid was minimaal. De raad had een sterke oriëntatie op de buitenwereld en had als belangrijkste ambitie het ziekenhuis een plaats te geven in de top 10 van de beste ziekenhuizen in de lijstjes die het Algemeen Dagblad en Elsevier jaarlijks uitbrengen. Aandacht voor het primaire proces was er niet of nauwelijks. Het primaire proces werd vooral overgelaten aan de directie. De aandacht voor bedrijfseconomische processen was doorgeslagen. Overigens bestond de Raad van Bestuur uit slechts één persoon. De commissie constateert dat de bestuurder van de organisatie in hoge mate verantwoordelijk is voor de ontstane organisatiecultuur. Over de Raad van Toezicht stelt de commissie dat de raad zich heeft laten verblinden door de stijging van het ziekenhuis op de scorelijsten van het Algemeen Dagblad en Elsevier.

Interessant in het rapport zijn ook de beschouwingen van de commissie over de historische achtergronden van het ziekenhuis. Na een periode van ernstige financiële zorgen is het ziekenhuis onder de betrokken bestuurder er weer bovenop gekomen en heeft het ziekenhuis een voortrekkersrol vervuld bij de introductie van marktwerking in de zorg in 2006. Ziekenhuizen moesten met elkaar concurreren om contracten met zorgaanbieders en om de gunst van patiënten. Zorginstellingen werden zelf na de afschaffing van het zogenaamde bouwregime, verantwoordelijk voor de kosten van renovatie en nieuwbouw. Het ziekenhuis heeft getracht de eigen concurrentiepositie te vergroten door veel aandacht te besteden aan financiële processen en vooral voor marketing en promotie. Centraal in de sturing van de directie stonden de prestatie-indicatoren op de Balanced Scorecard (van Kaplan en Norton), een ook in Nederlandse publieke organisaties veelvuldig gehanteerd instrument.

Het rapport van de commissie Lemstra roept vele vragen op die niet enkel betrekking hebben op het Maasstad Ziekenhuis. De meest fundamentele vraag betreft de marktwerking zelf. Vormt het functioneren van het Maasstad Ziekenhuis een uitzondering of is er sprake van een structurele fout in de vormgeving van het zorgstelsel? Kortom, ligt het aan de marktwerking of aan het Maasstad Ziekenhuis? Het is een vraag die uiteraard niet door de commissie kan worden beantwoord, maar die wel prangend naar voren komt. Een andere meer fundamentele vraag betreft de structuur van zorginstellingen, waarbij er sprake is van een vrij autonome en machtige medische staf (of beter: individuele medische professionals die veelal zelf niet in dienst zijn bij het ziekenhuis) die zich niet of nauwelijks laten aansturen, terwijl de Raad van

Bestuur van het ziekenhuis wel verantwoordelijk is voor de integrale kwaliteitszorg. Ook dit betreft een vraag die niet door de commissie kon worden beantwoord, maar die wel duidelijk naar voren komt.

De meest concrete vraag die naar voren komt en waar de commissie wel op ingaat, betreft de rol van de scorelijstjes van Elsevier en het Algemeen Dagblad. De score van het ziekenhuis blijkt centraal te hebben gestaan in de strategie van de bestuurder en de Raad van Toezicht van het Maasstad Ziekenhuis. Volgens het Algemeen Dagblad behoorde het ziekenhuis tot de beste van ons land.

Uit het rapport van de commissie wordt duidelijk dat er desondanks het nodige schortte aan de kwaliteit van het ziekenhuis (met name de primaire processen in het algemeen en de veiligheid in het bijzonder), maar dat dit gebrek aan kwaliteit kennelijk niet of nauwelijks verdisconteerd werd in de lijstjes van Elsevier en het Algemeen Dagblad. Het lijkt dan ook geen toeval dat juist in dit ziekenhuis de uitbraak zo slecht werd aangepakt.

Methode van onderzoek en analyse

In ons onderzoek hebben wij, mede naar aanleiding van de perikelen rondom het Maasstad Ziekenhuis, getracht een analyse te maken van de wijze waarop met name Elsevier en het Algemeen Dagblad tot hun oordelen komen. In methodische termen kan worden gesproken van een secundaire analyse van data. Opvallend hierbij was dat het nog niet zo gemakkelijk was te achterhalen hoe beide bladen tot hun oordelen zijn gekomen. Achterliggende stukken op websites boden wel enige uitkomst, maar van een grote mate van transparantie kan niet worden gesproken. Daarbij hebben wij bovendien opvallende telfouten geconstateerd. Een dergelijke secundaire analyse achten wij van groot maatschappelijk belang gezien de belangrijke rol die de lijstjes van Elsevier, het Algemeen Dagblad en de gegevens van de Inspectie Volksgezondheid spelen in de strategie en het beleid van zorginstellingen, verzekeraars, het ministerie, maar ook bij de keuze die patiënten maken bij het kiezen van een ziekenhuis (wanneer zij daartoe althans in staat zijn). Met de verwachte toenemende marktwerking zal het belang van deze lijstjes alleen maar groter worden.

De ‘beste ziekenhuizen’ (van Elsevier en het Algemeen Dagblad)

Wie opgenomen wordt in een ziekenhuis wil uiteraard de beste zorg. De trend in de gezondheidszorg is dat de patiënt steeds meer keuzemogelijkheden krijgt. Daartoe worden ook steeds meer de prestaties geobjectiveerd en openbaar gemaakt. Mede op basis van deze gegevens publiceren het Algemeen Dagblad en Elsevier jaarlijks hun ranglijsten met ‘beste’ en ‘slechtste’ ziekenhuizen. Vorig jaar publiceerden ze hun bijdrage over het jaar 2011. Uit een vergelijking vallen de enorme verschillen op. De drie beste ziekenhuizen volgens Elsevier, het Sint Antonius Ziekenhuis in Nieuwegein, het Meander Medisch Centrum in Amersfoort en het Albert Schweitzer Ziekenhuis in Dordrecht, staan daarentegen in het Algemeen Dagblad slechts op de plaatsen 5, 18 en 54 (van 93 ziekenhuizen). Naast de algemene score, op basis

van naar eigen zeggen objectieve data, heeft het Algemeen Dagblad ook onderzoek gedaan naar patiënttevredenheid over de deskundigheid van de artsen, de kwaliteit van de verzorging en van het ontslagproces en nazorg. De drie beste ziekenhuizen volgens Elsevier, scoren hierop in de middenmoot op de 31e, de 53e en de 35e plaats. Wat is hier aan de hand?

De drang naar transparantie en het nemen van onderbouwde beslissingen heeft een interessant fenomeen opgeleverd: lijsten. Rankings, als fenomeen ‘overgewaaid’ uit de Verenigde Staten, worden steeds meer gemeengoed in Nederland. Voor veel keuzes kunnen lijsten worden geraadpleegd. Zo zijn er rankings voor (middelbare) scholen, universiteiten, terrassen, en kunnen er minstens een keer per jaar oliebol- en haringtesten teruggevonden worden in de dagbladen. En zo geldt dit ook voor ziekenhuizen. Belangrijk voor de consument van zorg, zo lijkt.

Toegegeven, hoewel er over smaak niet te twisten valt, er kan voor enkelvoudige producten als haring en oliebol een oordeel worden gegeven over de kwaliteit. Dit geldt des te meer omdat de bestanddelen worden getest door een extern panel. Bij samengestelde producten of diensten als medische zorg ligt het een stuk gecompliceerder. Dit blijkt als de ziekenhuisrankings van het Algemeen Dagblad en Elsevier nader onder de loep worden genomen. Bovendien is maar net de vraag wie de gegevens aanlevert, en wat dit betekent voor de waarde van die gegevens? Leerden we niet allemaal op onze (top?)scholen: *Garbage in is garbage out*. Oftewel, de waarde van de resultaten hangt (mede) af van de betrouwbaarheid van de data.

In dit artikel wordt aandacht besteed aan de verschillen, de overeenkomsten en de conclusies van de ziekenhuislijsten die door het Algemeen Dagblad en Elsevier zijn opgesteld. De performance bubble wordt doorgeprikt: het beste ziekenhuis bestaat niet.

De criteria

Wat direct opvalt is dat er tussen de lijsten verschillen zijn in criteria voor kwaliteit. Er bestaat ook geen vastgestelde receptuur voor kwaliteit. En daar zit nu juist een van de knelpunten: wat bepaalt de kwaliteit? In de VS wordt bijvoorbeeld door US News bij de kwaliteit meegewogen of de specialisten het ziekenhuis aanmerken als een voor een bepaalde aandoening goede instelling. De reputatie weegt mee: *A survey of physicians, who are asked to name hospitals they consider tops in their specialty, produces a reputation score that is also factored in* (bron: US News 2011).

De criteria die door het Algemeen Dagblad en door Elsevier zijn gehanteerd, worden door beide toegelicht op hun website. Hieronder wordt beschreven waar naar eigen zeggen op beoordeeld is.

Het Algemeen Dagblad werkt met 38 kwaliteitscriteria. Hiervan zijn 33 criteria afkomstig van de Inspectie voor de Gezondheidszorg, of van medische werkgroepen van het project Zichtbare Zorg, waarin de Orde van Medisch Specialisten, zorgverzekeraars, patiëntenorganisaties en verenigingen van medisch specialisten samenwerken. Deze 33 criteria hebben betrekking op medische handelingen, bijvoorbeeld het aantal her-operaties, de aanpak van ondervoeding en het aantal risicovolle operaties. Criteria die voor iedere patiënt van belang zijn, als de kwa-

liteit van de verpleging, zijn ook meegewogen. Tevens is er overleg geweest met medische en brancheorganisaties om te bepalen in hoeverre een criterium iets zegt over de kwaliteit van zorg en om te bepalen of er duidelijke medische normen bestaan waaraan een ziekenhuis zou moeten voldoen. Vier van de 38 gebruikte criteria zijn gericht op patiëntvriendelijkheid en patiënttevredenheid. Twee betreffen patiëntenkeurmerken die zijn vastgesteld en toegekend door patiëntenverenigingen.

Elsevier vroeg aan bijna 5000 deskundigen wat het beste ziekenhuis van Nederland is. In tegenstelling tot de eerder genoemde Amerikaanse publicatie worden hier ook bestuurders en managers onder gerekend. Hoe hoger het percentage positieve respondenten, hoe meer punten zijn toebedeeld op zeven hoofdpunten: het medisch functioneren, de kwaliteit van de specialismen, de faciliteiten, de kwaliteit van onderzoek en opleiding, de kwaliteit van de verpleging, het personeel en de organisatie, omgang met de patiënt en de samenwerking met andere zorginstellingen en huisartsen.

Kortom, de ranking wordt in belangrijke mate bepaald door een verschillende opvatting over het begrip kwaliteit en de wijze waarop de gegevens worden verkregen en door wie ze worden aangeleverd.

Betrouwbaarheid en strategisch gedrag

Uit het proefschrift van Wim van de Bosch blijkt dat niet de kwaliteit van zorg, maar vooral meetfouten de gestandaardiseerde ziekenhuissterftecijfers beïnvloeden. Het door het Algemeen Dagblad meegenomen onderdeel sterftecijfer voor ‘cardiologie’ is dus naar alle waarschijnlijkheid onbetrouwbaar. Daar komt nog bij dat de scores fouten bevatten. Bij het onderdeel ‘Gebroken heup’ waar maximaal twee punten kunnen worden toebedeeld, scoren diverse ziekenhuizen echter drie punten. En zo zijn in elk geval tien fouten vast te stellen bij het Algemeen Dagblad.

In haar proefschrift concludeert Fleur Hasaart dat de financiering van ons zorgstelsel strategisch gedrag uitlokt bij artsen en zorginstellingen om hun inkomen op te krikken. Ook bij de ranking van ziekenhuizen ligt strategisch gedrag op de loer. Zo is een criterium dat in beide ranglijsten een rol speelt of er jaargesprekken met de medewerkers zijn geweest. De vraag rijst onmiddellijk wat de inhoud van deze gesprekken is geweest en wat er met de uitkomsten gebeurt, laat staan of ze werkelijk hebben plaatsgevonden. Grote stijgingen in de Algemeen Dagblad-lijst, zoals UMC St. Radboud Nijmegen en het Havenziekenhuis Rotterdam van respectievelijk plaats 75 naar 10 en 81 naar 20, zijn al te mooi. Dermate grote wijzigingen in de kwaliteit impliceren een enorme cultuuromslag die in een jaar tijd niet waar te maken is.

Belangrijk in de beoordeling is het aanleveren van volledige gegevens. Doelverschuiving kan optreden: het volledig registreren van alle handelingen kan het doel worden in plaats van goede medische zorg. Daarbij bestaat het risico dat er handelingen worden verricht die niet per se nodig zijn maar gedaan worden omdat dit ‘scoort’. Uit een publicatie in Medisch Contact (Nr. 03-16 januari 2007) blijkt dat sommige instellingen met definities sjoemelen: “Sommige instellingen draaien de inspectie en de media een rad voor ogen. Ze bestempelen

doorligplekken als vochtigheidsletsel en pronken zo met teruglopende decubituscijfers.” Het meest eerlijke ziekenhuis scoort dus wellicht veel lager op de ranglijst dan het daadwerkelijk zou moeten scoren.

Een belangrijke wetmatigheid binnen prestatiemeting is dat naarmate er meer belang aan wordt gehecht, de perverse effecten groter zijn. Zeker nu (directeuren van) ziekenhuizen zelf de gegevens aanleveren waarop ze worden beoordeeld, bestaat het gevaar dat er met gegevens wordt ‘gespeeld’, er staat immers veel op het spel, zeker ook voor de bestuurders van deze instellingen. De slager keurt zijn eigen vlees!

Het invoeren van prestatiemeting, kent daarom behalve voordelen wellicht nog meer nadelen. Een van de nadelen die door Hans de Bruijn (2006) wordt genoemd, heeft hij als wetmatigheid getypeerd: “De wet van de afnemende effectiviteit”. Hoe harder je afrekent op prestaties, des te meer perverse effecten zullen er optreden. De Bruijn gaat in zijn boek uitgebreid in op de perverse effecten van prestatiemeting in publieke organisaties. Overigens wordt uit de kredietcrisis overduidelijk dat perverse effecten van prestatiemeting en –beloning zich ook in private organisaties voordoen (de bonussen van onder andere de bankiers). Ook uit de eerder aangehaalde literatuur over de gezondheidszorg komt een soortgelijk beeld naar voren. Uit het in 2011 verschenen proefschrift van Van de Bosch blijkt dat niet de kwaliteit van zorg, maar vooral meetfouten de gestandaardiseerde ziekenhuissterftecijfers beïnvloeden. In haar recent verschenen proefschrift stelt Hasaart (2011) dat de financiering van ons zorgstelsel strategisch gedrag uitlokt bij artsen en zorginstellingen om zo hun inkomen op te krikken. Uit het rapport blijkt, zoals hierboven al eerder aangegeven, dat de Raad van Toezicht zich heeft laten verblinden door de stijging van het ziekenhuis op de scorelijsten van het Algemeen Dagblad en de Elsevier. De aandacht voor kwaliteit en veiligheid was minimaal. Dit terwijl de lijstjes volgens de opstellers onder andere de kwaliteit en veiligheid juist wel zouden meten. Uit bovenstaande analyse komt ook naar voren dat ‘kwaliteit’ geen eenduidige maatstaf is, maar –nog los van de betrouwbaarheid van de gegevens– op sterk uiteenlopende wijze ingevuld en geoperationaliseerd kan worden.

Conclusie

Het verdient aanbeveling dat het rapport van de commissie Lemstra een vervolg krijgt. Het is niet aannemelijk dat de negatieve effecten van prestatiemeting en het effect van scorelijstjes van het Algemeen Dagblad en de Elsevier zich beperken tot het Maasstad Ziekenhuis. Soortgelijke effecten zouden zich ook in andere ziekenhuizen (en wellicht breder: andere zorginstellingen) voor kunnen doen.

De kwalificatie van het beste of slechtste ziekenhuis zoals beide bladen ronkend weten te melden, doet geen recht aan de realiteit maar creëert een schijnwerkelijkheid. De gegevens zijn deels onbetrouwbaar, de meetmethodes zijn aanvechtbaar en de eindoordelen zijn arbitrair. Dit alles maakt dat de eindoordelen zelfs ten dele nietszeggend zijn. Het zou Elsevier en het Algemeen Dagblad sieren terughoudender te zijn bijvoorbeeld door het achterwege laten van totaaloordelen (Elsevier) en het opstellen van een top 100 (Algemeen Dagblad). Ziekenhuizen zijn nu eenmaal geen oliebollen of haringen.

Wil prestatiemeting echt zin hebben dan moeten ook de geleverde data door een onafhankelijk panel van deskundigen aan een kwalitatieve toets worden onderworpen. Zeker nu er steeds meer belang wordt gehecht aan ranglijsten, er beleid wordt gemaakt op door ziekenhuizen zelf geleverde data en er zelfs in toenemende mate financiële consequenties aan verbonden worden, is dit broodnodig. Prestatiemeting blijft anders niet meer dan een zeepbel waar teveel waarde aan wordt gehecht. ■

Referenties

- Bosch, W. van de, De HSMR beproefd: Aard en invloed van meetfouten bepalen van het gestandaardiseerde ziekenhuissterftecijfer, Amsterdam, 2011.
- Bovens, M.A.P., Verantwoordelijkheid en organisatie, beschouwingen over aansprakelijkheid, institutioneel burgerschap en ambtelijke ongehoorzaamheid, Zwolle, 1990.
- Bruijn, H. de, Prestatiemeting in de publieke sector, Den Haag, 2006.
- Commissie Lemstra (externe onderzoekscommissie MSZ), Oog voor het onzichtbare, Onderzoek naar de uitbraak van de Klebsiella Oxa-48 bacterie in het Maasstad Ziekenhuis Rotterdam, Rotterdam, maart 2012.
- Hasaart, F., Incentives in the diagnosis treatment combination payment system for specialist medical care, Maastricht, 2011.

Abstract

Some time ago the Commission Lemstra presented its report on the Maasstad Hospital Rotterdam, where the klebsiella bacterium prevailed. What stands out, in the report is the dominant role that the lists (best hospitals) of Elsevier and the Algemeen Dagblad play in the strategy of this hospital (and probably many others). This article focuses on the differences and similarities between, and the conclusions drawn from these rankings. The performance bubble is punctured: the best hospital does not exist. Firstly, it is clear that there is no single definition of quality employed; the research is based on varying characteristics. Additionally, data are supplied by the hospitals or hospital directors themselves, which generates strategic behavior and the reliability of the data is negatively affected. Thirdly, errors appear in the assigned scores. In short, the data are partly unreliable, the measurement methods are questionable and the final judgments are arbitrary.

