

Reputatiestrijd tussen ziekenhuizen komt kwaliteit niet ten goede

De ranglijst is een slechte raadgever

Heleen Pons, student
Zorgmanagement, instituut
Beleid & Management Gezond-
heidszorg (iBMG)

Hester Lingsma, promovendus,
afdeling Maatschappelijke
Gezondheidszorg, Erasmus MC

Roland Bal, hoogleraar bestuur
en beleid gezondheidszorg,
iBMG

Correspondentieadres:
r.bal@bmg.eur.nl;
c.c.: redactie@medischcontact.nl

Geen belangenverstrengeling
gemeld.

Ziekenhuisranglijsten kunnen reputaties maken en breken. De waarde ervan blijkt echter zeer relatief. Ziekenhuizen moeten zich dus niet gek laten maken en kunnen hun energie beter steken in een goed intern kwaliteitssysteem.

Oktober 2004: precies op de dag van de kick off-bijeenkomst van Sneller Beter, het nationale kwaliteitsprogramma voor de ziekenhuiszorg, wordt de eerste AD-ranglijst van ziekenhuizen gepubliceerd. Terwijl binnen kwaliteitsmanagers, adviseurs en vertegenwoordigers van de deelnemende ziekenhuizen luisteren naar speeches over de grote ambities van het programma, staat een aantal ziekenhuisbestuurders de hele dag op de gang te bellen. Met de lokale pers om uitleg te geven over de (slechte) score die hun ziekenhuis heeft behaald, met de voorlichtingsafdelingen om de marketingstrategie te bepalen. In één klap wordt duidelijk dat kwaliteit een kwetsbaar product is, waarbij externe beoordelingen reputaties op het spel kan zetten.¹

Reputatiemarkt

In korte tijd hebben de ziekenhuisranglijsten furore gemaakt. De 'markt' van de gezondheidszorg lijkt vooral een reputatiemarkt te zijn geworden. Voorlichtingsdiensten zijn inmiddels veranderd in marketingafdelingen, scores op prestatie-indicatoren en ranglijsten worden nauwlettend gevolgd en uitgebuit in de marketingstrategie. Reputatiemanagement heeft zijn intrede gedaan in de ziekenhuiswereld.

Maar waar de ranglijsten van AD, Elsevier en Roland Berger, en de keuzesites van KiesBeter en Independer reputaties kunnen maken en breken, komt ook hun eigen reputatie op het spel te staan.² In een onderzoek op aangeven van de NVZ stelde het MOA Center for Marketing Intelligence & Research onlangs vast dat de Elsevier-lijst onbetrouwbaar is.³ De op basis van *peer review* vastgestelde lijst zou manipuleerbaar zijn doordat ziekenhuizen hun specialisten (kunnen) oproepen het eigen ziekenhuis stelselmatig hoger te waarderen. Een aantal ziekenhuizen heeft op basis hiervan besloten niet langer deel te nemen aan het Elsevier-onderzoek.

Maar hoe zit het nu feitelijk met de betrouwbaarheid van de ranglijsten? Om hier achter te komen, hebben wij de ranglijsten van het AD en Elsevier van 2005 tot en met 2008 met elkaar vergeleken. Keert de 'puntenwolk van Giard' die in eerder onderzoek werd beschreven hier terug of is de onderlinge vergelijkbaarheid inmiddels verbeterd?⁴ En is er nog steeds sprake van een 'jojo-effect' van ziekenhuizen door opeenvolgende jaargangen van de lijsten?⁵

Elsevier en AD

Elsevier was in 1996 de eerste die een ziekenhuisranglijst publiceerde. De gegevens waarop Elsevier de ranglijsten baseert, zijn door de jaren gelijk gebleven: specialisten, huisartsen, hoofdverpleegkundigen, kwaliteitsmanagers en raden van bestuur beoordelen de ziekenhuizen op medische kwaliteit door middel van

Specialisten kunnen het eigen ziekenhuis stelselmatig hoger waarderen



Ziekenhuizen jojoën door de opeenvolgende jaargangen van de lijsten

een enquête. De beoordeling per ziekenhuis bestaat uit een puntentelling, gebaseerd op wat de deskundigen als sterk of juist als zwak punt aankruisen.

De gegevens waarop het AD de ranglijst baseert wijken hier sterk vanaf. Het AD maakt namelijk gebruik van openbare kwaliteitsgegevens. Voor de ziekenhuisranglijst van 2008 waren dat bijvoorbeeld 26 geselecteerde kwaliteitsaspecten waarop de ziekenhuizen punten kunnen scoren. Daarnaast worden wel patiëntenoordelen, verkregen van Independer, in de ranglijst meegenomen.

Geen correlatie

Al onze analyses zijn gebaseerd op 73 ziekenhuizen, waarvan op alle lijsten gegevens beschikbaar waren. Om de onderlinge vergelijkbaarheid van de ranglijsten te bepalen, zijn

Spearman correlatiecoëfficiënten uitgerekend, zowel tussen de verschillende jaargangen van één ranglijst als tussen verschillende ranglijsten.

Het blijkt dat alleen de correlaties tussen de verschillende jaargangen van dezelfde lijst enige samenhang vertonen. Er is geen correlatie tussen de ranglijsten van het AD en die van Elsevier.

Om te kijken of eerdere jaargangen van de ranglijsten voorspellende waarde hebben voor de jaren daarna, is een multivariate lineaire regressie uitgevoerd. Daaruit bleek dat voor zowel het AD als Elsevier alleen het jaar 2007 enige voorspellende waarde heeft voor 2008; daarvoor is voorspelling op basis van de ranglijstpositie niet mogelijk. De gegevens uit 2007 zijn echter ook niet erg voorspellend voor 2008. Ook hebben we de onderlinge voorspellende waarde tussen de verschillende ranglijsten onderzocht. Hier werd niets significant gevonden. Het AD is dus niet voorspellend voor het volgende jaar van Elsevier of andersom. Verder is gekeken naar de trend in de tijd van enkele individuele ziekenhuizen. Hiervoor zijn door toeval 7 van de 73 ziekenhuizen geselecteerd en de plaatsen van deze ziekenhuizen op de ranglijsten zijn uitgezet in de tijd. In de grafiek is te zien dat er een enorme fluctuatie bestaat in de opeenvolgende jaren. Voor de meeste ziekenhuizen is geen consistent patroon te ontdekken in de tijd.

Elsevier iets beter

We moeten concluderen dat we opnieuw te maken hebben met de eerdergenoemde puntenwolk van Giard. Er zijn veel ziekenhuizen die bij het AD hoog op de ranglijst staan maar bij Elsevier juist laag, en andersom. Dit is ook te verwachten aangezien Elsevier en het AD andere dingen meten. Kwaliteit is niet eenduidig en dat komt in de ranglijsten ook tot uiting. Maar is er ook een beste ranglijst en zo ja, welke? De ranglijst van Elsevier staat ter discussie vanwege het gebruik van *peer review*. Juist omdat collega's oordelen moeten geven over andere ziekenhuizen zou de lijst niet betrouwbaar zijn. Dat kennelijke wantrouwen in *peer review* is vreemd. Dit systeem staat in allerlei andere domeinen juist centraal, bijvoorbeeld in de wetenschap, waar collega's elkaars publicaties en subsidieaanvragen beoordelen. Hoewel ook dit problematisch kan zijn, blijft *peer review* het dominante systeem.⁶ Elsevier gebruikt het adagium dat 'dokters onderling prima weten wat het goede ziekenhuis is' en dat lijkt zo gek nog niet.

SAMENVATTING

- Met de publicatie van ranglijsten van ziekenhuizen staan zowel de reputaties van ziekenhuizen als die van de ranglijsten zelf inmiddels op het spel.
- Statistische analyse van de Elsevier- en AD-lijsten laat zien dat de lijsten onderling niet vergelijkbaar zijn.
- De lijst van Elsevier doet het iets beter waar het gaat om de consistentie door de jaren heen.
- Geen van beide lijsten is betrouwbaar. Ziekenhuizen doen er goed aan zich te richten op hun interne kwaliteitssystemen.

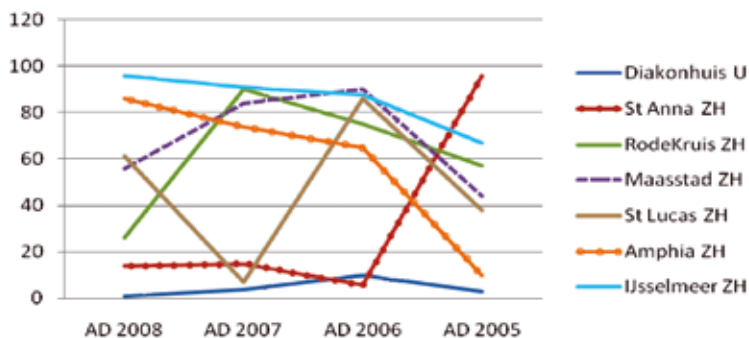
Het AD gebruikt de prestatie-indicatoren van IGZ. Hoewel prestatie-indicatoren intuïtief objectiever lijken dan opvattingen van collega's is hiervan echter bekend dat de objectiviteit problematisch is en dat de statistische betrouwbaarheid vaak zeer laag is als gevolg van kleine patiëntenaantallen per ziekenhuis.⁷ Aan beide methoden kleven dus nadelen. Bij gebrek aan een gouden standaard is het dan ook niet mogelijk te zeggen welke van de lijsten

het beste is. Belangrijker lijkt daarom de consistentie in de tijd, en wat dat betreft doet het *peer review*-systeem van Elsevier het iets beter dan het systeem gebaseerd op prestatie-indicatoren van het AD. Dat ziekenhuizen van positie op de ranglijst veranderen in opeenvolgende jaren, is op zich goed. Daarmee laten de lijsten zien dat ze gevoelig zijn voor verandering in de kwaliteit van zorg. Het zou dan echter te verwachten zijn dat die verandering van posities zelf consistent zou zijn. Dat is niet het geval: individuele ziekenhuizen jojoën nog steeds door de opeenvolgende jaargangen van de lijsten. Dit betekent dat het voor het publiek (patiënten, verzekeraars, ziekenhuizen, overheid) niet mogelijk is 'goede' en 'slechte' ziekenhuizen te identificeren op basis van de Elsevier- noch de AD-lijst, hoewel als gezegd de Elsevier-lijst hier iets beter scoort.

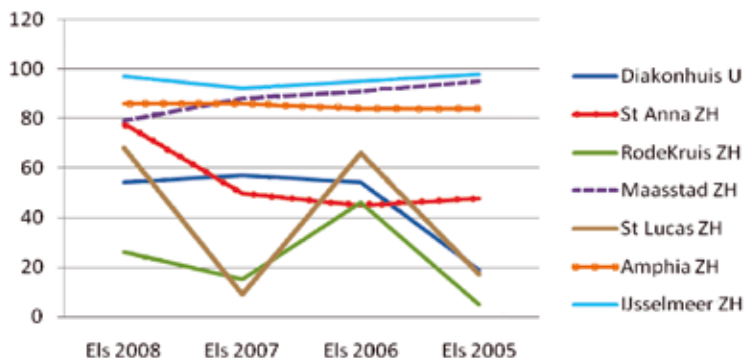
Ranglijsten

De plaatsen van zeven at random geselecteerde ziekenhuizen op de ranglijsten van het AD (boven) en Elsevier (onder) door de jaren heen. Voor de meeste ziekenhuizen is geen consistent patroon te ontdekken.

Trend AD-ranglijst



Trend Elsevier-ranglijst



Aura van objectiviteit

Ondanks het gebrek aan betrouwbaarheid, spelen de lijsten een belangrijke rol en richten ziekenhuizen hun gedrag erop in.⁸ Met name de AD-lijst, met het aura gebaseerd te zijn op objectieve prestaties, lijkt gevolgd te worden. Dit leidt tot steeds meer nadruk op reputatiemanagement, maar gaat dus blijkbaar niet vanzelfsprekend gepaard met kwaliteitsverbetering. Dit verschijnsel is al eerder aangetoond in de financiële sector en de wetenschap.² MOA Center for Marketing Intelligence & Research en eerder de Raad voor de Gezondheidszorg hebben opgeroepen tot een sectorbrede score.⁹ Een dergelijke score kan wellicht helpen de huidige verwarring over ranglijsten te voorkomen, maar het is niet erg waarschijnlijk dat die er gaat komen (de reputatiemarkt kent zijn eigen dynamiek), noch dat die een objectieve graadmeter gaat vormen (want gegevens blijven afkomstig van instellingen die op die reputatiemarkt concurreren). Verdergaande validering zal bovendien een groeiende registratielast met zich meebrengen. De strijd op de reputatiemarkt zal dus nog wel doorgaan. Maar intussen doen ziekenhuizen en andere zorginstellingen er goed aan zich te richten op kwaliteitsverbetering op basis van hun interne kwaliteitssystemen. 



De bijbehorende literatuurlijst en meer MC-artikelen over ranglijsten vindt u bij dit artikel op onze website: www.medischcontact.nl.

Referenties

- 1 Zuiderent-Jerak, Teun. 2009. Competition in the wild. Emerging figurations of healthcare markets. *Social Studies of Science*, 39: 765-792.
- 2 Power, Michael, Tobias Scheytt, Kim Soin, and Kerstin Sahlin. 2009. Reputational risk as a logic of organizing in late modernity. *Organization Studies* 30 (2 & 3):301-324.
- 3 van Hamersveld, Mario, and Lex Olivier. 2009. *Ziekenhuiskenners Keuren Zichzelf*. De Elsevier Onderzoeksaanpak Doorgelicht. Amsterdam: MOA, Center for Marketing Intelligence & Research.
- 4 Giard, R.W.M. 2005. Prestatie-indicatoren als maat voor kwaliteit van medische zorg: retoriek en realiteit. *Nederlands Tijdschrift voor Geneeskunde* 149 (49):2715-2719.
- 5 Bal, Roland. 2008. *De nieuwe zichtbaarheid*. Sturing in tijden van marktwerking, Oratiereeks Erasmus MC. Rotterdam: Erasmus MC.
- 6 Weingart, Peter. 2005. Impact of bibliometrics upon the science system: inadvertent consequences? *Scientometrics* 62 (1):117-131.
- 7 Bevan, Gwyn, and Christopher Hood. 2006. What's measured is what matters. Targets and gaming in the English public health care system. *Public Administration* 84 (3):517-538.
- 8 van Woensel, Marco, Pieter van der Valk, and Leo te Velde. 2007. Effecten van ziekenhuisranglijsten. *ZM magazine* 23 (9):2-4.
- 9 http://www.rvz.net/cgi-bin/nieuws.pl?niew_srcID=228