

Segnalare ai cittadini la qualità degli ospedali: il rischio della cattiva informazione

Public reporting on individual hospitals' quality: the risk of misinformation

Gian Luca Di Tanna, Luca Cisbani, Roberto Grilli

Area governo clinico, Agenzia sanitaria regionale Emilia-Romagna

Corrispondenza: Gian Luca Di Tanna, Agenzia sanitaria regionale Emilia-Romagna, viale Aldo Moro 21, 40127 Bologna; e-mail: GDiTanna@Regione.Emilia-Romagna.it

Cosa si sapeva già

■ L'attenzione dedicata dai media alla performance delle strutture sanitarie e alla qualità dei servizi erogati è sempre maggiore. Esiste una crescente consapevolezza relativa non solo alle implicazioni delle iniziative di pubblicizzazione dei risultati dei servizi sanitari, ma anche alla complessità tecnico-metodologica del rappresentare le performance cliniche in modo adeguato, accurato e di facile comprensione.

Cosa si aggiunge di nuovo

■ Questo articolo, sulla base di due casi ritenuti «esemplari», documenta come il modo in cui viene affrontato il problema della comunicazione al pubblico di informazioni sulla qualità dei servizi sia ancora largamente inadeguato.

Riassunto

Obiettivo: benché con sempre maggiore frequenza i mass media riportino informazioni sulla qualità dei servizi, spesso con l'esplicito intento di individuare i «migliori», la comunicazione di questi temi avviene attraverso modalità che o sono troppo semplicistiche nel loro sforzo di confrontare le performance cliniche dei singoli centri, o ricorrono a indicatori di qualità inadeguati. In tale senso, queste informazioni possono facilmente portare a interpretazioni errate. Questo lavoro considera due esempi concreti per evidenziare il problema.

Metodi: sono stati utilizzati due casi recenti di comunicazione attraverso i mass media delle performance cliniche di centri cardiocirurgici: i risultati dello studio BPAC «Esiti a breve termine di interventi di by-pass coronarico nelle cardiocirurgie italiane», coordinato dall'ISS, e la pubblicazione sulla rivista

Panorama di quelle che erano definite come le migliori cardiocirurgie italiane sulla base di un indice di reputazione.

Risultati: il primo dei casi citati evidenzia i problemi della rappresentazione delle performance con semplici *league tables*. Con un metodo basato su Markov Chain Monte Carlo si evince che il ricorso alle classifiche, pur basate su un corretto indicatore di esito, è un esercizio potenzialmente fuorviante. Nel secondo caso, si dimostra l'inaffidabilità di un indice reputazionale utilizzato come indicatore della qualità dei servizi.

Conclusione: le modalità con cui si comunica al pubblico la qualità dei servizi sono ancora largamente inadeguate e quindi potenzialmente fuorvianti nell'indirizzare le scelte dei cittadini. (*Epidemiol Prev* 2006; 30(3): 199-204)

Parole chiave: *league tables*, monitoraggio delle performance cliniche, indice di reputazione, interventi di by-pass, metodi bayesiani

Abstract

Objective: increasingly information on individual hospitals' clinical performance is disclosed to the general public through lay mass media. Usually the explicit goal of those initiatives is to guide citizens' choice identifying the «best» centres. However, these efforts rely either on simplistic approaches in comparing the clinical performance of individual hospitals or on inadequate quality indicators. Therefore, this information may easily misguide the general public. In this paper, in order to highlight the problems implied, we consider two recent cases of mass media reporting on hospitals' quality.

Methods: two recent cases of mass media reporting on cardiac surgery centres' quality are considered: the BPAC study «Esiti a breve termine di interventi di by-pass coronarico nelle cardiocirurgie italiane» coordinated by the Italian Institute of Health, and the publication by the newsmagazine *Panorama* of the list of the «best» Italian cardiac surgery centres based on a reputational index.

Results: the first example points out the problems implied by presenting hospitals' clinical performance through league tables. Using the Markov Chain Monte Carlo method, this study shows how the conclusions merely based on ranking – even when these are based on a sensible quality indicator – can be potentially misleading. As for the second case study, it is shown the inadequacy of the quality indicator (the reputation index) adopted.

Conclusion: despite the growing awareness in the health services community of the technical problems implied by describing and comparing the quality of care provided by individual hospitals, «real life» examples of provision of information on quality of care by the mass media to the general public appear to be largely unsatisfactory and potentially misleading. (*Epidemiol Prev* 2006; 30(3): 199-204)

Keywords: *league tables*, clinical performance monitoring, reputational index, by-pass procedures, bayesian methods

Introduzione

Il fatto che i mezzi di comunicazione di massa guardino con attenzione alla sanità non rappresenta certo una novità. I temi che attengono alla salute sono ovviamente al centro dell'attenzione dei cittadini anche, e forse soprattutto, nelle implicazioni relative al funzionamento dei servizi sanitari. Da questo punto di vista, l'attenzione ai temi sanitari rappresenta per i mass media un mezzo per catturare il pubblico.

Tuttavia, la relazione tra mass media e servizi sanitari negli ultimi anni è stata certamente enfatizzata da alcuni fenomeni, tra loro connessi, che hanno caratterizzato la sanità a livello internazionale. Ci riferiamo sia al crescente ricorso ai mass media da parte di organismi istituzionali come strumento per veicolare ai cittadini informazioni sulla qualità dei servizi, sia alla rilevanza acquisita dal tema della scelta degli utenti in ambito sanitario. Relativamente al primo aspetto, la divulgazione delle performance comparative dei servizi (solitamente nella forma di graduatorie, *league tables*)¹⁻⁵ rappresenta un'opzione sempre più spesso considerata come strumento per enfatizzare la responsabilizzazione degli operatori, per stimolare un'attenzione costante alla qualità delle prestazioni nell'ambito delle organizzazioni sanitarie e, infine, per promuovere l'*empowerment* dei cittadini come attori in grado di operare scelte consapevoli in ambito sanitario, decidendo per esempio a quale ospedale/professionista rivolgersi per farsi curare. E' esattamente su questo punto che le politiche di pubblicizzazione dei risultati clinici dei servizi si interfacciano con il tema più generale delle scelte di cittadini e pazienti, viste come strumento per promuovere una maggiore flessibilità da parte dei servizi nei confronti delle preferenze dei singoli come delle comunità e, in buona sostanza, per attivare in ambito sanitario dinamiche non dissimili da quelle che caratterizzano i mercati, viste come foriere di maggiore efficacia ed efficienza.

L'effettiva resa di queste scelte di politica sanitaria è per la verità alquanto controversa. Tuttavia, in questa sede vogliamo focalizzare la nostra attenzione principalmente sui problemi posti dal ruolo dei mass media come mezzi attraverso i quali comunicare ai cittadini la qualità dei servizi e su come l'informazione potrebbe essere meglio veicolata.

In particolare, utilizzando due casi italiani recenti, intendiamo portare l'attenzione sui problemi che emergono sia quando i media sono semplicemente lo strumento che veicola informazioni che altri (tipicamente istituzioni di ricerca o di go-

verno) raccolgono sulle performance cliniche degli ospedali, sia quando i media stessi sono promotori di inchieste finalizzate a documentare la qualità dei servizi sanitari, in tutti e due i casi avendo come presumibile obiettivo quello di promuovere scelte informate da parte dei cittadini.

Materiali e metodi

I caso: lo studio sulla mortalità a 30 giorni per by-pass aorto-coronarico

Nel febbraio 2005 l'Istituto superiore di sanità ha reso pubblici sul web (sito internet: <http://bpac.iss.it>) i risultati di uno studio condotto su 81 cardiocirurgie italiane e finalizzato a documentare la mortalità a breve termine per interventi di by-pass aorto-coronarico isolato. In particolare, i risultati erano pubblicati con una modalità che consentiva l'individuazione del nominativo del singolo centro, della sua mortalità osservata e di quella attesa sulla base delle procedure di aggiustamento mediante standardizzazione indiretta. In questo modo, era facilmente derivabile una classifica che consentiva di individuare i centri «migliori» e quelli «peggiori» (figura 1). Questo è appunto ciò che alcuni mezzi di comunicazione di massa hanno fatto.⁶⁻⁹

Ai fini di questo lavoro abbiamo voluto, senza mettere in alcun modo in discussione l'approccio adottato dall'Istituto superiore di sanità (ISS) sul piano metodologico, analizzare le implicazioni della semplice presentazione di graduatorie che non tengano conto del grado di incertezza insito nell'attribuzione di una specifica «posizione» (*rank*) in classifica a un singolo centro.^{10,11} In particolare, i dati resi disponibili dall'ISS sono stati rianalizzati mostrando gli intervalli di credibilità dei *rank* assegnati a ogni centro, introducendo l'elemento «errore di campionamento» tramite l'utilizzo della tecnica Markov Chain Monte Carlo. Infatti, il metodo «classico» per confrontare istituzioni differenti consiste nel disegnare il grafico dei tassi osservati (meglio se *risk-adjusted*) e il loro intervallo di confidenza (solitamente al 95%). Se l'intervallo non si sovrappone a quello definito come *benchmark* (il valore medio nazionale in questo caso), il centro è definito «critico» (non necessariamente nell'accezione negativa del termine). Così facendo ci si può aspettare di identificare il 2,5% dei centri come significativamente al di sotto (o al di sopra) dello standard per il solo effetto del caso, anche se in realtà negli stessi si riscontra un livello di performance in linea con quello medio. Questo è il motivo per cui

Tabella 1. Tassi di mortalità aggiustati per interventi di by-pass isolato dei centri di cardiocirurgia in base al corrispondente indice di reputazione.

Table 1. Case-mix adjusted mortality rates for isolated coronary artery by-pass graft of cardiac surgery centres, according to their index of reputation.

Indice di reputazione	n. ospedali	RAMR medio	DS	range		n. ospedali sopra il RAMR medio nazionale (2,61%)
più basso	6	2,28	1,94	0,84	6,05	1
intermedio/basso	7	2,53	1,32	1,15	5,30	2
intermedio/alto	7	2,85	1,44	1,24	5,25	2
più alto	6	3,04	2,26	0,36	7,20	3

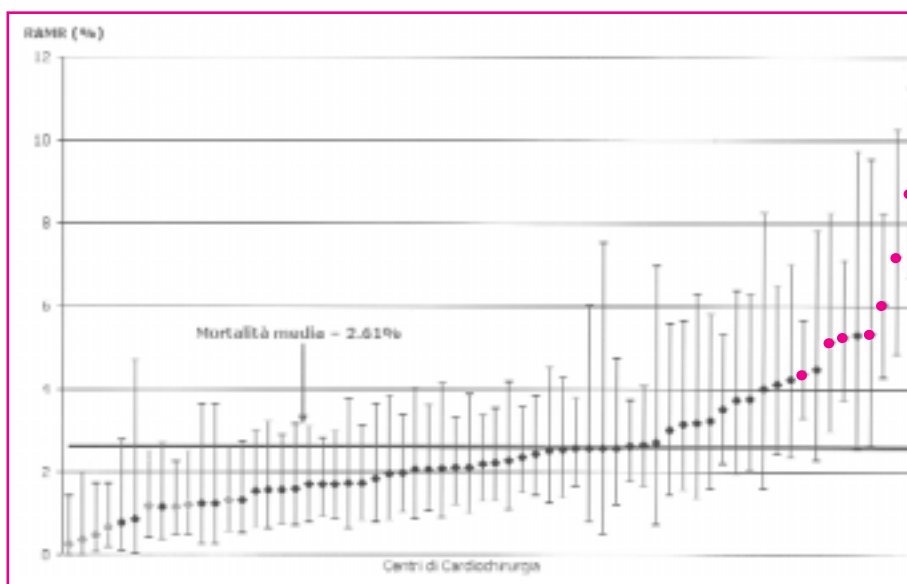


Figura 1. Risk Adjusted Mortality Rate (RAMR) per centro (con IC 95%). I pallini in grigio segnalano i centri statisticamente significativi al di sotto della media, in colore i centri statisticamente significativi al di sopra.

Figure 1. Risk Adjusted Mortality Rate (RAMR) for hospital (with 95% CI). Grey dots show the hospitals statistically significant below the national average, colored ones the hospitals statistically significant above it.

la pubblicazione e l'interpretazione dei risultati «statisticamente significativi» va fatta con estrema cautela, anche perché con questa procedura in realtà non si fa che testare l'ipotesi che ogni centro abbia esattamente lo stesso tasso di mortalità, assunzione né plausibile, né interessante. In ogni caso, sono stati individuati 8 centri «migliori» e 7 «peggiori» (figura 1).

A partire dai dati pubblicati, abbiamo quindi effettuato alcune simulazioni associando a ciascun centro una distribuzione di probabilità (centrata sul tasso di mortalità riscontrato nel centro stesso e alla quale può essere data un'interpretazione bayesiana) e da questa è stato estratto un tasso «vero»: ottenuto così un insieme di tassi, questi possono essere classificati e tale procedura reiterata fino a convergenza delle stime (ritenuta soddisfacente quando l'errore Monte Carlo di ogni parametro di interesse è risultato inferiore ad almeno il 5% della deviazione standard). E' possibile quindi ricavare per ogni centro una distribuzione dei posti in classifica raggiunti.

Sono stati utilizzati diversi modelli (con l'uso del software WinBUGS, disponibile all'indirizzo internet <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml>) a partire da differenti distribuzioni di probabilità (Binomiale, Poisson) con l'utilizzo o meno di *prior* informative (con valore di riferimento pari a quello nazionale) e considerando sia l'ipotesi di indipendenza dei centri (modello a effetti fissi) sia quella di scambiabilità (modello a effetti casuali) con tassi veri (su scala logit) estratti da una comune distribuzione normale.

Il caso: le classifiche delle cardiocirurgie basate sulla loro reputazione

Da settembre 2005 la rivista *Panorama* ha pubblicato i risultati di un'indagine condotta su un campione di 2.047 medici di medicina generale.¹² Questi ultimi erano stati invitati a segnalare, per diverse aree specialistiche, a quali centri avrebbero avviato i propri pazienti in caso di necessità. Sulla base

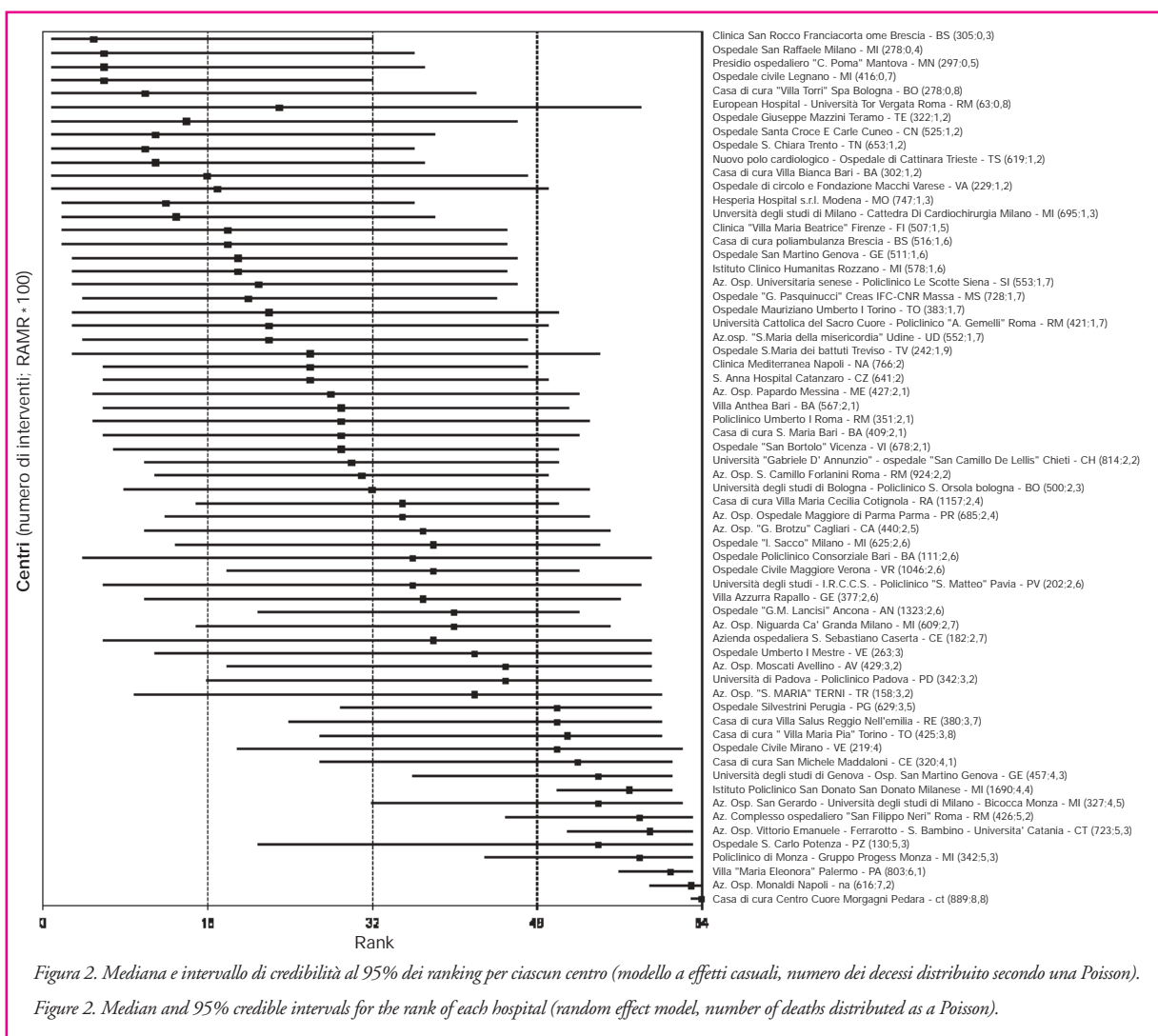
del numero di volte in cui uno specifico ospedale era citato dai medici intervistati è stato costruito un indice di reputazione e, sulla base di questo parametro, sono stati identificati centri qualificati come «di eccellenza».

In particolare, sul numero di *Panorama* del 16.09.2005 sono stati pubblicati i risultati relativi alle cardiocirurgie, dieci delle quali erano poste in particolare evidenza per aver avuto un elevato indice di reputazione.

In questo caso specifico, traendo vantaggio dalla disponibilità dei dati relativi al già citato studio sulla mortalità per by-pass aorto-coronarico, abbiamo cercato di verificare se e in che misura la mortalità per questo specifico intervento chirurgico avesse una qualche relazione con l'indice di reputazione utilizzato per descrivere la qualità dei centri cardiocirurgici. In breve, utilizzando i dati riportati sul sito web <http://bpac.iss.it> per 26 centri cardiocirurgici (quelli ai quali era associato uno score di «reputazione») è stato possibile studiare la relazione tra la mortalità aggiustata a 30 giorni per by-pass aorto-coronarico e il loro score reputazionale riportato dalla rivista. In aggiunta a una semplice regressione tra queste due variabili, i centri sono stati raggruppati in classi in base alla distribuzione in quartili dell'indice di reputazione e per ciascun quartile è stata stimata la mortalità media. Ovviamente, l'ipotesi di partenza era che il quartile con l'indice di reputazione più elevato fosse quello con la minore mortalità cardiocirurgica.

Risultati

I caso. La figura 2 illustra i risultati delle nostre analisi sulla *league table* dei centri cardiocirurgici basata sulla mortalità a 30 giorni per by-pass aorto-coronarico. I risultati in termini di *rank* mediani (più alto il *rank*, peggiore il risultato) e relativi intervalli di credibilità non differiscono a seconda del modello considerato (il grafico presenta il *ranking* in base al mo-



dello Poisson a effetti casuali con distribuzioni *a priori* non informative uniformi). Gli intervalli sono molto ampi, mostrando una grande incertezza associata ai posti in classifica ottenuti. I 15 centri considerati «outlier» poiché statisticamente significativi possono essere posizionati nella migliore o peggiore metà: nessun centro può essere collocato nel «migliore» quarto, mentre 5 centri risultano essere nel quarto «peggiore» e meritano quindi un approfondimento.

In sostanza, questa modalità di presentazione dei risultati consente di evidenziare con ragionevole certezza che vi sono alcuni centri con performance cliniche peggiori degli altri, mentre risulta molto più difficile identificare i cosiddetti «migliori».

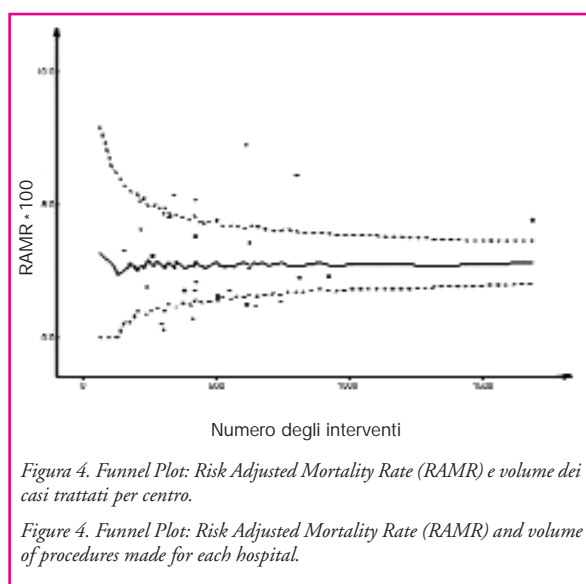
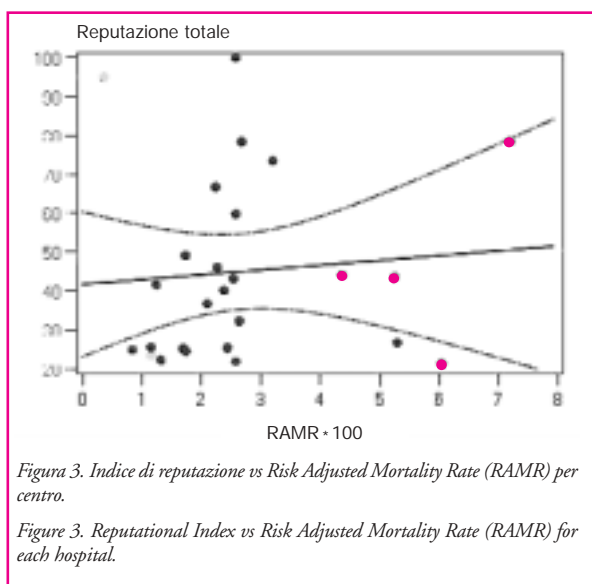
Il caso. Come mostrato in tabella 1, i centri che ricadono nel primo quartile (ovvero quelli con l'indice reputazionale più basso) sono quelli che in media presentano il RAMR (Risk Adjusted Mortality Rate) più basso. Tra i «migliori» 6 centri

in base alla loro reputazione, 3 hanno il RAMR più alto rispetto alla media nazionale (2,61%). Nel complesso, l'indice reputazionale non è correlato con la mortalità (coefficiente di correlazione di Pearson= 0,087, $p= 0,674$) (figura 3).

Solo uno dei centri nei primi dieci della classifica stilata in base alla reputazione ha un tasso di mortalità significativamente più basso rispetto a quello atteso (aggiustato per il *case-mix*); viceversa, sono due i centri che, seppure considerati tra i migliori in base alle opinioni dei medici coinvolti nella ricerca, presentano una mortalità significativamente più elevata di quella attesa.

Discussione e conclusioni

I due casi citati sono a nostro avviso sufficientemente rappresentativi delle diverse tipologie di iniziative finalizzate a comunicare ai cittadini la qualità dei servizi sanitari, avendo l'obiettivo di orientarne le scelte individuando i centri «migliori» e «peggiori».



Nel primo caso, la pubblicazione di classifiche dei centri cardiocirurgici basata sulla mortalità per by-pass aorto-coronario isolato ha indotto i mass media all'identificazione di centri qualificati come migliori, sulla base di una modalità di presentazione (le *league tables*) semplicistica.¹³ Analisi più sofisticate, come si è visto, hanno portato a concludere che le performance di questi centri non differivano sostanzialmente dalle performance degli altri.

Nel secondo caso, per certi aspetti ancora più eclatante, la classifica pubblicata era basata su un indicatore di qualità (l'indice di reputazione) che si è dimostrato non correlato con l'effettiva qualità tecnica delle prestazioni erogate dai singoli centri, quest'ultima rappresentata dalla mortalità a 30 giorni dall'intervento.

I due casi citati non costituiscono in alcun modo esempi isolati, anzi sono rappresentativi di un problema di carattere generale già emerso nell'ambito della letteratura internazionale,¹⁴ problema destinato a diventare tanto più rilevante quanto più si diffondono le iniziative di pubblicizzazione dei risultati dei servizi e finanche dei singoli operatori.¹⁵⁻¹⁸ Nel contesto di questo lavoro non abbiamo inteso entrare nel merito dei numerosi aspetti problematici che simili iniziative sollevano sul piano delle relazioni tra servizi sanitari e cittadini, tra i professionisti clinici e le organizzazioni sanitarie nel contesto delle quali operano, e infine tra medici e pazienti. In questo senso, ci siamo volutamente soffermati su alcuni aspetti tecnici, piuttosto che sulle implicazioni di *policy*. Tuttavia vale la pena sottolineare, a prescindere da quale sia l'opinione di ciascuno relativamente alla controversa opportunità di rendere pubbliche ai cittadini informazioni sulle performance cliniche, come le esigenze politiche tendano a fare aggio ai problemi di metodo, con risultati potenzialmente negativi. Da questo punto di vista, il recente caso del National Service Sco-

tland (NHS),¹⁹ «costretto» a rendere pubbliche statistiche di mortalità ospedale-specifiche nonostante siano basate su dati amministrativi inaffidabili, rappresenta un evidente esempio di come l'imperativo di informare sempre e comunque rischi di portare alla circolazione di dati fuorvianti o errati.

Porre quindi l'accento su alcune questioni di metodo non significa in alcun modo dimenticare che gli aspetti più rilevanti della pubblicizzazione dei risultati sono rappresentati dalle implicazioni politiche in senso lato. Piuttosto, rappresenta un richiamo alla necessità di tenere adeguatamente conto degli aspetti metodologici nella loro realizzazione pratica.

Sul versante dei metodi vi sono almeno tre diversi aspetti da considerare, due dei quali compresi nella nostra valutazione: la scelta degli indicatori utilizzati per descrivere la qualità dei servizi, le fonti informative su cui la loro rilevazione è basata e, infine, le modalità di rappresentazione dei risultati.

Cominciando da quest'ultimo aspetto, nonostante ne sia messa in discussione la validità,^{5,13,17,20,21} l'uso delle classifiche appare essere ancora la modalità di rappresentazione dominante. Certamente questo approccio ha l'ovvio vantaggio di essere semplice da rappresentare e apparentemente di facile comprensione. Tuttavia, come si è cercato di dimostrare anche in questo lavoro mediante l'utilizzo delle MCMC in un'ottica bayesiana, rappresenta un'opzione facilmente fuorviante.

D'altronde, quanto più sofisticate e complesse diventano le analisi, tanto più difficile è rappresentarle in modo chiaro i risultati: in questo senso, uno dei temi centrali della ricerca in questo settore è trovare un compromesso equilibrato tra validità scientifica e comprensibilità.

Tra le soluzioni possibili, il Funnel Plot (figura 4) rappresenta un valido strumento di analisi e di comunicazione dei risultati.²² Nato come strumento utilizzato nelle metanalisi delle revisioni sistematiche per individuare l'eventuale presenza

di *publication bias*, sempre più spesso viene utilizzato anche in analisi di valutazione delle performance in sanità. L'indicatore di mortalità è rappresentato in base alla misura della sua precisione: la numerosità dei casi osservati e i limiti di controllo creano una sorta di «imbuto» intorno all'esito atteso. Nel nostro caso, il grafico rivela come gran parte dei centri siano all'interno dei limiti al 95%, determinando tuttavia l'individuazione di alcuni centri «critici».

Gli altri due aspetti di metodo problematici riguardano, come si è detto, la scelta degli indicatori e le fonti utilizzate per la loro rilevazione.

La gran parte delle iniziative di pubblicizzazione dei risultati sono fortemente incentrate sull'analisi della mortalità ospedaliera. Questo non è sorprendente, dal momento che si tratta di un indicatore relativamente facile da rilevare e di forte impatto emotivo sui destinatari dell'informazione. In questo senso non è casuale che l'ambito cardiocirurgico sia tra quelli più frequentemente indagati da iniziative di pubblicizzazione, non solo per la sua ovvia rilevanza e visibilità, ma anche per la relativa facilità di rilevazione degli esiti di interesse. Tuttavia, su un piano più generale è evidente come vi sia il rischio che le comparazioni tra servizi finiscano per concentrare la propria attenzione non su quello che è importante conoscere, in quanto rilevante per gli operatori e i pazienti, ma su ciò che è appunto facilmente misurabile.

Peraltro, la spinta alla ricerca di indicatori semplici è probabilmente la spiegazione del ricorso, come nel caso citato in questo articolo, a un indicatore quale l'indice di reputazione, dimostratosi totalmente inaffidabile come descrittore della qualità dei servizi.

Infine, la fonte di rilevazione vede il contrasto tra l'esigenza di ricorrere a fonti di informazione facilmente disponibili, quali i database amministrativi, e quella di avere a disposizione informazioni sufficientemente dettagliate sulle caratteristiche e la complessità dei singoli pazienti assistiti nei centri oggetto della valutazione. Questo per le ovvie implicazioni sulla possibilità di evidenziare quanto le performance cliniche osservate siano condizionate non dalla capacità degli operatori e dei servizi, ma dalle caratteristiche prognostiche dei pazienti. Anche in relazione a questo specifico aspetto non vi sono soluzioni semplici al problema. Se il citato studio coordinato dall'ISS sul by-pass aorto-coronarico ha rappresentato un buon esempio di come questo problema possa essere risolto attraverso una rilevazione *ad hoc*, basata sull'attivo coinvolgimento dei centri partecipanti, in altre circostanze si è dimostrato che i database amministrativi, adeguatamente integrati da alcune informazioni rilevanti, siano altrettanto validi di più complesse (e costose) rilevazioni *ad hoc*.²³

In conclusione, la pubblicizzazione dei risultati dei servizi attraverso i mass media non dovrebbe prescindere dalla considerazione di aspetti metodologici rilevanti ai fini di garantire l'attendibilità delle informazioni veicolate. Attualmente,

al di fuori della ristretta cerchia degli addetti ai lavori in questo ambito della ricerca sanitaria, il livello di consapevolezza sulle problematiche in esame appare ancora clamorosamente scarso.

Conflitti di interesse: nessuno

Ringraziamenti. Un ringraziamento particolare al Professor David Spiegelhalter per la cortese disponibilità e i preziosi suggerimenti.

Bibliografia

- Jacobs R, Smith PC. *A descriptive analysis of general acute Trust star ratings*. Centre for Health Economics, University of York, York (UK), 2004. Discussion Paper.
- Keogh B, Spiegelhalter D, Bailey A, Roxburgh J, Magee P, Hilton C. The legacy of Bristol: public disclosure of individual surgeons' results. *BMJ* 2004; 329, 450-454.
- Rowan K, Harrison D, Brady A, Black N. Hospitals' star ratings and clinical outcomes: ecological study. *BMJ* 2004; 328, 924-925.
- Hannan EL, Vaughan-Sarrazin MS, Doran DR, Rosenthal GE. Provider profiling and quality improvement efforts in coronary artery bypass graft surgery. The effect on short-term mortality among Medicare beneficiaries. *Med Care* 2003; 41(10), 1164-1172.
- Jacobson B, Mindell J, McKee M. Hospital mortality league tables. *BMJ* 2003; 326: 777.
- Cardiocirurgia, solo 8 i centri eccellenti. *Il Messaggero*, 08.02.2005.
- Perrone M. Cuore, ecco gli ospedali migliori. *Il Sole 24 Ore*, 08.02.2005.
- Satolli R. Bypass, i migliori. *Corriere Salute*, 20.02.2005.
- Satolli R. Bypass, i centri migliori dove operarsi. *Corriere della Sera*, 09.02.2005.
- Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. *Health Technology Assessment* 2000; 4(38), 1-130.
- Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Methods in health service research: An introduction to bayesian methods in health technology assessment. *BMJ* 1999; 319: 508-512.
- Grandi Ospedali – Una guida per curarsi meglio: Cardiocirurgia, cardiologia. *Panorama*, 29.09.2005.
- Adab P, Rouse AM, Mohammed MA, Marshall T. Performance league tables: the NHS deserves better. *BMJ* 2002; 324: 95-98.
- Krumholz HM, Rathore SS, Chen J, Wang Y, Radford MJ. Evaluation of a consumer-oriented internet health care report card. The risk of quality ratings based on mortality data. *JAMA* 2002; 287: 1277-1287.
- Bridgewater B. Why doctors' outcomes should be published in the press. *BMJ* 2005; 331: 1210.
- Marshall MN, Romano PS. Impact of reporting hospital performance. *Quality and Safety in Health Care* 2005; 14: 77-78.
- Shahian DM, Normand SL, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE et al. Cardiac surgery report cards: comprehensive review and statistical critique. *Ann Thorac Surg* 2001; 72: 2155-2168.
- Marshall M, Shekelle P, Brook R, Leatherman S. *Dying to know: Public release of information about quality of health care*. Nuffield Trust and RAND, London (UK), 2000.
- Christie B. Scottish health service is ordered to release data on surgeons' performance. *BMJ* 2005; 331: 1424-d.
- Bird SM, Cox D, Farewell VT, Goldstein H, Holt T, Smith PC. Performance indicators: good, bad, and ugly. *J R Stat Soc* 2005; 168 part 1: 1-27.
- Marshall T, Mohammed MA, and Rouse A. A randomized controlled trial of league tables and control charts as aids to health service decision-making. *Int J Qual Health Care* 2004; 16(4), 309-315.
- Spiegelhalter DJ. Funnel plots for institutional comparison. *Qual Saf Health Care* 2002; 11: 390-391.
- Fortuna D, Vizioli M, Contini A, Guidoni I, Biagi B, Pacini D et al. Assessing clinical performance in cardiac surgery. Does a specialised clinical database make a difference? *Interact CardioVasc Thorac Surg* 2006; 1-12.