

RECENTI SVILUPPI NELLA METODOLOGIA STATISTICA PER LA VALUTAZIONE DELL'EFFICACIA DEGLI OSPEDALI

G. Vittadini
M. Sanarico
C. Rossi

INTRODUZIONE

È sorta negli ultimi anni l'esigenza di valutare la qualità di strutture operanti nei settori dei servizi di pubblica utilità alla persona quali istruzione, sanità, assistenza etc.

Alle analisi di efficienza che considerano il rapporto tra risorse impiegate e risultato di gestione (Aigner, Lovel, Schmidt, 1977; Gori, Rossi, Grassetti, 2001), ai modelli per lo studio della soddisfazione dell'utente (Parasuraman ed al., 1985; Zanella 2000; D'Ambra 2001), si sono aggiunti studi di efficacia relativa (Goldstein e Spiegelhalter, 1996). In tali studi, mediante il modello multilevel, si costruiscono graduatorie di strutture sanitarie¹ misurate sui pazienti dopo l'erogazione del servizio, tenendo conto delle loro condizioni prima della cura². Nel presente lavoro si analizzano criticamente i più recenti sviluppi di questa metodologia con particolare attenzione alla sua applicazione al caso delle strutture ospedaliere. Un primo filone di sviluppo riguarda le scelte metodologiche, nell'ambito del modello multilevel, rispetto alla natura e alla distribuzione degli indicatori osservati e degli outcome, alle variabili casuali che individuano l'efficacia degli ospedali, al modello nel suo complesso. Un secondo filone riguarda il dibattito sulla scelta degli outcome e degli indicatori atti a misurare rispettivamente l'efficacia e le condizioni dei pazienti prima della cura, a partire da dati di origine clinico-epidemiologici, osservazionale, amministrativa. Un'indagine empirica riguardante il complesso dei ricoverati in area chirurgica degli ospedali lombardi nel 2001 esemplifica i risvolti operativi ed innovativi di quanto presentato in chiave teorica.

¹ Intesa come "la capacità di un intervento sanitario di modificare la condizione del paziente portandola o riportandola a uno stato migliore del precedente" fino ad avvicinarsi alla "effectiveness" (efficacia praticamente raggiungibile) o all'*efficacy* (efficacia attesa massima). (Donabedian 1990, Pagano e Rossi, 1999, pag. 267).

² Altrimenti le strutture ospedaliere per non peggiorare la loro valutazione, possono esercitare una "selezione avversa" volta ad escludere dal ricovero i pazienti più gravi (Goldstein, Spiegelhalter, 1996).

1. SCELTE METODOLOGICHE NEGLI STUDI DI EFFICACIA RELATIVA DEGLI OSPEDALI

Nel lavoro di Goldstein e Spiegelhalter (1996), punto di riferimento sotto il profilo metodologico, il modello *multilevel* (Hox, 1995) è utilizzato per studiare l'efficacia relativa gli ospedali. Si ha quindi per l' i -esimo paziente curato nell' j -esimo ospedale:

$$y_{ij} = \alpha_q x_{ijq} + u_j + e_{ij} \quad (i=1, \dots, n; \quad j=1, \dots, g, \quad q=1, \dots, p) \quad (1)$$

ove y_{ij} è l'outcome sanitario; x_{ijq} sono le variabili esplicative che descrivono le caratteristiche dei pazienti, in particolare il loro stato di salute prima della cura; α_q sono i parametri ad esse connessi; u_j è una variabile casuale che individua l'efficacia relativa del j -esimo ospedale al netto delle caratteristiche dei pazienti; e_{ij} è il residuo casuale. Si ipotizza che i residui e_{ij} siano mutualmente incorrelati con valore atteso nullo e varianza costante e che ogni variabile casuale u_j abbia distribuzione $\varphi(u_j)$ con varianza costante³.

Il modello, denominato "a effetti misti", permette di valutare l'efficacia relativa perché, grazie alla struttura gerarchica dei dati, enuclea la variabilità tra ed entro gli ospedali, *coeteris paribus*, al netto delle caratteristiche dei pazienti, superando nello stesso tempo i problemi di mancanza di informazione dovuti a piccoli campioni grazie all'introduzione dei "shrinkage residuals"⁴ (Goldstein e Spiegelhalter, 1996, Marshall e Spiegelhalter, 2001 p.128). Le soluzioni sono ottenute mediante metodi di inferenza bayesiana (con Software Bugs Gibbs Sampling basato su un particolare algoritmo Monte Carlo Markov Chain). Proposta perciò una distribuzione a priori $\varphi_1(u_j)$ per la variabile casuale u_j , si ottiene la distribuzione a posteriori $\varphi_2(u_j)$ a partire dalla quale si costruiscono graduatorie di efficacia tra ospedali.⁵ Fermo restando l'uso del modello

³ Può essere intesa come distribuzione dell'*outcome* al netto delle caratteristiche individuale o "risk adjusted", vale a dire ottenuta mediante procedura di "risk adjustment" (Thomas, Longford, Rolph, 1994; JCHAO 2000).

⁴ Una volta ricavato il valore atteso $E(u_j)$ del parametro casuale u_j lo si corregge mediante il

fattore *shrinkage* nel seguente modo: $\hat{u}_j = \frac{n_j s_a^2}{n_j s_a^2 + s_e^2} E(u_j) \quad [E(u_j) \neq 0]$

ove s_a^2 e s_e^2 sono rispettivamente la varianza fra e negli ospedali. Quanto più piccole sono la numerosità del sotto campione j -esimo e la varianza negli ospedali tanto è maggiore la correzione per $E(u_j)$. (Healy 2001, p 8)

⁵ Si assume che un ospedale A sia migliore di un altro ospedale B se l'estremo inferiore dell'intervallo di confidenza ricavato dalla distribuzione $\varphi_2(A)$ è superiore all'estremo superiore dell'intervallo di confidenza ottenuto dalla distribuzione $\varphi_2(B)$. Alternativamente, estendendo il criterio proposto da Dagum (1987) per l'indice del REA (*Relative Economic Affluence*), si possono confrontare i valori attesi di $\varphi_2(A)$ e di $\varphi_2(B)$. La posizione nella graduatoria di ogni ospedale sarà determinata dal numero di confronti con esito positivo con ogni altro ospedale.

multilevel, negli anni intercorsi tra il lavoro di Goldstein e Spiegelhalter ⁶ (1996) e il momento attuale sono state avanzate proposte di generalizzazione sotto il profilo metodologico, in alcuni casi accompagnate da studi empirici, a riguardo di tre argomenti:

a) Quando l'outcome è identificato da un solo indicatore discreto, la sua distribuzione è usualmente binomiale (o poissoniana) nell'ambito di modelli *logit* o *probit*; se invece l'indicatore è continuo di solito l'indicatore è normale (Davis e Gibbens, 1995; Morris e Christiansen, 1996; Scott e Shiell, 1997; Rice, 2001; Langford e Day, 2001; Marshall e Spiegelhalter, 2001). Numerose sono però le proposte alternative⁷ anche in situazioni in cui vi siano outliers (Smans 1992; Langford e Lewis 1988) o dati mancanti tra gli outcome (Leyland e Boddy, 1998).

Qualora gli *outcome* siano misurati mediante una pluralità di indicatori si propongono modelli di regressione lineari multivariati multilevel con errori incorrelati (risolvibili con il metodo dei minimi quadrati)⁸ oppure, qualora si voglia tener conto esplicitamente dei legami tra diversi indicatori, si utilizzano modelli multivariati con errori correlati quali i modelli Sure (Vittadini e Sanarico, 2003). Alternativamente, definito l'outcome come costruito teorico multidimensionale sottostante l'insieme di indicatori osservati z_{iju} ($u=1, \dots, d$) e da essi misurato a meno di errori δ_{iju} , lo si ricava sotto il profilo statistico come variabile latente y_{ij}^+ , mediante un modello fattoriale o con metodi a esso alternativi caratterizzati da soluzioni uniche quali il *Partial Least Squares* (Wold 1982):

$$z_{iju} = \mathbf{L}_{uj} y_{ij}^+ + \delta_{iju} \quad (i=1, \dots, n; j=1, \dots, g; u=1, \dots, d) \quad (2)$$

⁶ Nel lavoro di Goldenstein e Spiegelhalter (1996) dopo uno studio sull'efficacia relativa degli ospedali scozzesi nella cura dell'infarto miocardico basato sul tasso di mortalità non "*risk adjusted*", si propone uno studio dell'efficacia di differenti chirurghi nell'operazione di by-pass all'arteria coronaria. Il campione è composto da 87 pazienti in ospedali dello Stato di New York; l'outcome sanitario è il tasso di mortalità intra ospedaliero dopo l'operazione distribuito come una variabile casuale di Poisson; vi è una sola variabile esplicativa inerente la gravità dei pazienti prima dell'operazione espressa da punteggi di origine epidemiologica; gli \mathbf{u}_j sono relativi a differenti chirurghi; $\varphi_1(\mathbf{u}_j)$ e $\varphi_2(\mathbf{u}_j)$ sono distribuiti in modo normale.

⁷ Si propongono ad esempio outcome dicotomici dalla distribuzione non binomiale nell'ambito di modelli logit, come nel caso di studi sulla sanità pubblica in Gran Bretagna (Wagstaff e Van Doorslaer, 1993, O'Donnell, 1993); outcome normali di ordine p in uno studio sull'efficacia relativa di ospedali lombardi rispetto a patologie cronico-degenerative (Vittadini e Sanarico, 2003). Nell'ambito di studi di statistica medica si propongono anche outcome distribuiti in modo multinomiale nominale (Kunin ed al., 1990; Li, 1997) o multinomiale ordinale (Hedeker e Gibbons, 1994; Yang, 2001).

⁸ Come nel caso di Mc Culloch ed al. (1997) ove si investiga la variabilità tra ospedali della lunghezza della degenza e dell'eventuale nuovo ricovero d'urgenza per pazienti operati.

per poi porre y_{ij}^+ come variabile dipendente nel modello *Multilevel* (1) (Goldstein e Leyland, 2001, p.186). Quando si abbiano indicatori qualitativi, definiti su scale nominali, dicotomiche, politomiche, dati i gravi problemi di unicità delle soluzioni dei modelli fattoriali definiti su indicatori misti (Vittadini, 1999) si ricavano gli *outcome* quali *proxies* quantitative delle variabili latenti attraverso opportuni metodi di *Multidimensional Scaling* o di *Latent Trait Analysis* quale la *Rasch Analysis*⁹ (Wright e Masters, 1982; Bond e Fox, 2002). Tali proxies sono poi utilizzate come variabili dipendenti in (1). Infine l'efficacia relativa può essere studiata anche nel caso in cui gli outcome siano definiti sulle strutture ospedaliere nel loro complesso e non su singoli pazienti (Goldstein et al., 2000).

b) Come in Goldstein e Spiegelhalter (1996) nei principali studi finora proposti l'efficacia delle strutture ospedaliere è analizzata mediante variabili casuali la cui distribuzione è ricavata attraverso metodi di inferenza bayesiana (Thomas et. Al, 1994; Goldstein e Spiegelhalter, 1996; Morris e Christiansen, 1996; Rice e Leyland, 1996; Leyland e Boddy, 1998; Marshall e Spiegelhalter, 2001). Perciò, una volta ipotizzata una distribuzione a priori $\varphi_1(\mathbf{u}_j)$ per le variabili casuali \mathbf{u}_j si utilizzano anche le informazioni provenienti dal campione per giungere ad una definitiva e più informativa distribuzione a posteriori $\varphi_2(\mathbf{u}_j)$ ¹⁰.

Nella quasi totalità delle proposte è assunta distribuzione a priori normale per $\varphi_1(\mathbf{u}_j)$. Non mancano però forti critiche per questa scelta (Marshall e Spiegelhalter 2001, p.139-140), soprattutto nel caso di outliers (Langford e Lewis, 1988). Per questo è stato proposto, nell'ambito dell'inferenza bayesiana di sfruttare a pieno la flessibilità dell'approccio Monte Carlo Markov Chain per proporre distribuzioni a priori troncate o basate su modelli mistura o modelli che tengano conto di outliers (Albert e Chib 1997; Richardson e Green 1997; Marshall e Spiegelhalter 2001).

c) Nel suo complesso, il modello può basarsi su dati di tipo epidemiologico raccolti secondo disegni sperimentali tipici dei *clinical trials* o su dati osservazionali, di origine campionaria o amministrativa.¹¹

⁹ Questa tecnica è particolarmente appropriata nel caso di indicatori ordinali originati da questionari, prove, test, poiché in questi casi, la quantificazione viene effettuata tenendo conto della difficoltà dei questionari e dell'abilità delle persone (Vittadini 2000).

¹⁰ La metodologia d'inferenza bayesiana: fornisce stime più accurate e robuste grazie al "fattore *shrinkage*"; costruisce le graduatorie sulla base della probabilità di non superare certe soglie espresse in valore assoluto e non solo sulla base di confronti tra ospedali; valuta il grado di incertezza delle graduatorie e l'incremento di conoscenza ottenibile aumentando l'ampiezza campionaria; permette di compiere studi longitudinali distribuendo l'informazione in diversi intervalli temporali; presenta indubbi vantaggi sul piano analitico e computazionale se gli ospedali di cui occorre studiare l'efficacia relativa sono numerosi (Thomas, Longford, Ralph (1994); Morris e Christiansen (1996)).

¹¹ E' evidentemente sempre preferibile far riferimento a dati epidemiologici raccolti nell'ambito di un opportuno disegno sperimentale piuttosto che a dati osservazionali.

In secondo luogo, nel caso si ipotizzi che i parametri relativi alle caratteristiche individuali dei pazienti varino casualmente tra gli ospedali, in luogo del modello “a effetti misti” (1) si ha il modello a “effetti casuali”¹² (Thomas, Longford, Rolph 1994) :

$$y_{ij} = (\alpha_j + f_{jq}) x_{ijq} + \mathbf{u}_j + e_{ij} \quad (i=1, \dots, n; j=1, \dots, g; q=1 \dots p) \quad (3)$$

con α_j parametro fisso e f_{jq} variabile casuale.

In terzo luogo, qualora si ipotizzi che l'efficacia di ogni ospedale sia differente per sottogruppi di pazienti eterogenei rispetto a qualche loro caratteristica inerente la salute, in luogo del modello (1) si ha¹³:

$$y_{ij} = (\bar{\alpha}_q + \delta_{jr}) x_{ijq} + \mathbf{u}_j + e_{ij} \quad (i=1, \dots, n; j=1, \dots, g; q=1, \dots, p; r=1, \dots, s) \quad (4)$$

con $\bar{\alpha}_q$ effetto medio dell'ospedale j -esimo e δ_{jr} variabile casuale che misura l'effetto interazione fra l'ospedale j -esimo e l' r -esimo sottogruppo di pazienti curati in tale ospedale.

Il modello *multilevel* (1) è estendibile al caso gerarchico di tre o più livelli (es: reparti/ospedali/ASL)¹⁴. Si ha:

$$y_{ijl} = \alpha_q x_{ijq} + \theta_j + e_{ijl} \quad (i=1, \dots, n; j=1, \dots, g; l=1 \dots, m; q=1, \dots, p) \quad (5)$$

$$\theta_j = \lambda_j + v_{ij} \quad (6)$$

ove θ_j individua l'efficacia associata all' l -esimo reparto del j -esimo ospedale e λ_j l'efficacia dell'ospedale nel suo complesso.

In questo modo:

$$v_{ij} = \theta_j - \lambda_j \quad (7)$$

¹² Mentre le (1) variano tra ospedale ed ospedale per le variabili casuali u_j che fungono da intercette, le (3) variano anche per le variabili casuali f_{jq} che fungono da coefficienti angolari.

¹³ Ad esempio Duncan ed al. (1998) mostrano come l'outcome medio di un ospedale risulti dagli effetti dalle caratteristiche dei pazienti, dall'efficacia media dell'ospedale al netto delle caratteristiche dei pazienti, e dall'interazione tra l'effetto-ospedale e caratteristiche dei pazienti; Leyland e Boddy (1998) studiano l'efficacia di ospedali scozzesi rispetto alla cura dell'infarto miocardico ipotizzando effetti diversi su sottogruppi di pazienti individuabili in base alle comorbidità.

¹⁴ Ad esempio, gli stessi Marshall e Spiegelhalter (2001) riprendendo Goldstein e Spiegelhalter (1996), propongono di considerare dati gerarchici a livello di paziente, chirurgo, ospedale.

esprime l'efficacia differenziale dovuta al reparto l -esimo dell'ospedale j -esimo al netto dell'efficacia attribuibile all'ospedale (Marshall e Spiegelhalter, 2001) ¹⁵.

Infine, utilizzando le proprietà teoriche del modello multilevel (Hox 1995), si può tener conto nell'analisi di efficacia relativa, delle caratteristiche delle diverse strutture ospedaliere z_{kj} (Gori e Vittadini, 1999). Data l'equazione (1), si riesprime la variabile casuale u_j nel seguente modo:

$$u_j = v_k z_{kj} + w_j \quad (j = 1, \dots, g; k = 1, \dots, f) \quad (8)$$

con v_k parametro fisso a livello di sistema, z_{kj} k -esima caratteristica dell' j -esimo ospedale e w_j residuo che descrive l'efficacia dell'ospedale j -esimo non solo al netto delle caratteristiche individuali dei pazienti ma anche delle caratteristiche delle strutture ospedaliere¹⁶. A secondo del tipo di variabili inerenti gli ospedali z_{kj} utilizzate in (8) e della natura delle variabili dipendenti y_{ij} in (1) cambia l'obiettivo dell'analisi.

TABELLA 1
Natura delle variabili e tipo di analisi

VARIABILI DIPENDENTI y_i	VARIABILI ESPLICATIVE Pazienti x_q	VARIABILI ESPLICATIVE Ospedali z_k	SCOPO
Output	Caratteristiche pazienti	Input ospedali-reparti	Efficienza
Outcome sanitario	Caratteristiche pazienti	Output ospedali	Efficacia al netto di efficienza
Outcome sanitario	Caratteristiche pazienti	Input-output ospedali	Efficienza + Efficacia

Si possono infine introdurre parametri di interazione γ_{kq} fra le variabili esplicative inerenti i pazienti x_q e le variabili riguardanti gli ospedali z_k .

¹⁵ Vi sono anche proposte in cui si considerano strutture di osservazioni "cross classified" (Rasbash e Browne 2001).

¹⁶ Negli Usa vi sono banche dati inerenti caratteristiche di strutture sanitarie quali ad esempio l'*Hedis (The Health Plan Employer Data)* (Epstein 1995) e l'*Uniform Hospital Discharge Data Set* (Dubois et al. 1988) che forniscono le informazioni più disparate sugli ospedali (dimensioni, utilizzazione dei servizi, occupazione dei letti, percentuale dei pazienti per reparto, informazioni su aspetti finanziari e nel management, case mix (vedi nota 46), caratteristiche medie dei pazienti ivi curati (quali lunghezza della degenza, tasso di morte, età, tasso di ricovero, provenienza, tipologia del ricovero etc). Anche in altre situazioni come in Gran Bretagna o in Scozia sono fornite analoghe informazioni ad opera del "National Health Service" e del "The Information and Statistic Division of the National Health Service". Perciò numerosi studi usano in modelli multilevel variabili esplicative inerenti gli ospedali definite a livello aggregato (Dubois et al. (1988); Morris e Christiansen (1996)) o a livello di singolo paziente (Leyland e Boddy, 1998).

Si ha in luogo di (8)

$$u_j = v_k z_{kj} + \gamma_{kq} (z_{kj} \bar{x}_{jq}) + w_j \quad (j=1, \dots, g; q=1, \dots, p; k=1, \dots, f) \quad (9)$$

con \bar{x}_{jq} media della variabile esplicativa q -esima nell'ospedale j -esimo.

1.1. Outcome e variabili esplicative

Il dibattito sulle scelte metodologiche, sviluppabile soprattutto a partire da studi di nicchia di tipo epidemiologico inerenti particolari patologie, evita eccessive semplificazioni rispetto alla realtà nella formulazione del modello multilevel. Per individuare il “valore aggiunto” in termini di salute dopo l'erogazione del servizio è necessaria però anche una opportuna selezione degli outcome e delle variabili esplicative

In letteratura l'outcome più usato nella maggior parte degli studi è il tasso di mortalità intraospedaliero entro 30 giorni dalla dimissione (Duncan 1988; Thomas, Longford e Rolph 1994; Leyland e Boddy 1995; Normand ed al. 1995; Goldstein e Spiegelhalter 1996; Marshall e Spiegelhalter 1999; Marshall e Spiegelhalter 2001). Tuttavia molti autori ritengono che non si possa individuare l'efficacia utilizzando esclusivamente tale *outcome* (Dubois ed al. 1988; Epstein 1995; Marshall e Spiegelhalter 2001; Goldstein e Spiegelhalter 1996; Schneider e Epstein 1997)¹⁷. Sono stati proposti, alternativamente, altri outcome rilevati da studi epidemiologici effettuati su piccoli campioni di pazienti¹⁸ e più raramente outcome di carattere più generale più facilmente rilevabili.¹⁹

Per ciò che concerne lo stato di salute del paziente prima della cura, in alcuni studi sono proposti indicatori che dovrebbero individuare la gravità del paziente per ogni patologia. Anche in questo caso non mancano critiche²⁰ e sono stati

¹⁷ Addirittura è considerato a volte causa di dimissioni ingiustificate e premature favorite dagli ospedali per evitare che la mortalità verificatasi prima del 30° giorno influenzino negativamente la valutazione dell'ospedale (Goldstein e Spiegelhalter 1996).

¹⁸ Ad esempio Morris e Christiansen (1996) utilizzano in uno studio inerente trapianti di rene in ospedali americani le frequenze relative dei rigetti; in Italia Rizzi ed al. (1999) propongono in uno studio inerente il tasso di incidenza endemico delle infezioni chirurgiche le infezioni ospedaliere postoperatorie e in uno studio inerente le lesioni da decubito la variabile ordinale di Norton.

¹⁹ Ad esempio Rice e Leyland (1996) propongono di usare in uno studio sulla cura del neoplasma del colon la lunghezza della degenza.

²⁰ Goldstein e Spiegelhalter (1996) affermano che indicatori di gravità delle malattie possono non tener conto dell'influsso di variabili sociali e altri fattori; Epstein (1997) afferma che la pubblicazione di dati relativi a pazienti che dovevano sottoporsi a by pass coronarico ha portato a una sopravvalutazione del rischio di mortalità; Leyland e Boddy (1998) e Marshall e Spiegelhalter (1999) osservano come gli indicatori di gravità usualmente proposti non siano sufficienti a descrivere il differente stato di salute dei pazienti prima della cura. Anche l'indice di gravità

perciò proposti indicatori che considerano le condizioni generali di salute del paziente²¹. Ciò premesso per la scelta di indicatori e outcome, si possono avanzare alcune proposte di ordine generale.

In primo luogo si possono utilizzare indicatori di outcome di origine clinica o epidemiologica, specifici per ogni patologia, costruiti secondo criteri desunti dalla letteratura scientifica, rilevati ad hoc o da cartelle cliniche o da documentazione direttamente connessa alla cura. Ciò già avviene in alcuni paesi per il tasso di mortalità che viene raccolto sistematicamente per ospedale ed aree geografiche²², in qualche caso tenendo conto delle caratteristiche individuali dei pazienti sottoposti a cura grazie all'utilizzo di metodologie uguali o simili a quella qui presentata²³. Per ciò che concerne gli altri outcome negli USA, dove vi sono forti investimenti finanziari anche privati in sanità e vi è la presenza di strutture scientifiche pubblico-private atte a rilevare dati clinici-epidemiologici, non mancano studi e pubblicazioni a riguardo di criteri per individuare e raccogliere in modo sistematico e critico indicatori²⁴. Tuttavia tali raccolte di indicatori, oltre

proposto dal “*The Prospective Payment System Rand Quality and Care Study*” (vedi nota 23) è criticato in quanto spiega molto poco la variabilità della mortalità. (Jenks et al 1988)

²¹ Ad esempio Rice e Leyland (1996) in uno studio inerente la cura di tumori del colon suggeriscono di utilizzare età, sesso, tipo di ricovero, tipo di dimissione dopo il ricovero, tipo di trattamento ricevuto durante il ricovero; Leyland e Boddy (1998) in uno studio inerente il tasso di mortalità per infarto miocardico propongono di usare quali variabili individuali età, sesso, stato civile tipo di ricovero, diagnosi secondaria, eventuale operazione chirurgica precedente. Marshall e Spiegelhalter (1999) sottolineano la necessità di usare un vasto insieme di indicatori inerente diagnosi, sesso, età e gravità della malattia per superare le differenze dovute alle diverse condizioni del paziente prima delle cure.

²² Esistono raccolte di tassi di mortalità per ospedali quali quella dell’*“Health Care Finance Administrative”* negli Stati Uniti, (Daley ed al. 1988) quella del *“National Health Service Management Executive”* in Inghilterra (Goldstein e Spiegelhalter 1996) e dello *“Scottish System of Hospital Discharge”* in Scozia (Leyland e Boddy 1998).

²³ A questo scopo Leyland e Boddy (1998) hanno utilizzato il modello multilevel, con variabili esplicative individuali provenienti dal *“General Register Office for Scotland”*, quali età, sesso, stato civile, diagnosi principale e secondaria all'accettazione e alcune variabili relative ai diversi ospedali; Dubois ed al. (1988) a partire da un campione di 205.000 pazienti di 93 ospedali americani, hanno impiegato un modello di regressione multipla su dati aggregati inerenti tipo di diagnosi e di ricovero, aspetti demografici dei pazienti, caratteristiche degli ospedali; il *“Medical Mortality Predictor System”* ed il *“Prospective Payment System”*, mediante il modello multilevel, forniscono periodicamente, tassi di mortalità *“risk adjusted”* medi per ospedali a riguardo di 4 patologie (infarto miocardico, scompenso cardiaco, polmonite, stroke) (Thomas Longford e Rolph 1994). Le informazioni a riguardo della gravità prima della cura sono ricavate dal *“Prospective Payment System Rand Quality of Care Study”* basato su 60-80 variabili inerenti condizioni di patologie e comorbità per pazienti acuti e cronici” rilevate a partire da cartelle cliniche su un campione di pazienti di ospedali americani (Daley ed al. 1988)

²⁴ Si hanno ad esempio:

a) gli insiemi di indicatori della *Healthcare Cost and Utilization Project (HCUP)* dell’*Agency for Health Research and Quality dell’US Department of Health and Human Services (AHRQ 2003)*. Gli indicatori a livello di ospedale sono indicatori di processo (indicatori di volumi, di prestazione e di utilizzazione

a lasciare aperti alcuni problemi di ordine metodologico,²⁵ non hanno portato ad un unico sistema di valutazione della qualità.²⁶ L'intendimento attuale è piuttosto quello di validare, all'interno del processo di accreditamento dell'JCAHO, sistemi di indicatori ed outcome proposti dalle strutture ospedaliere.²⁷ Tuttavia in generale e, soprattutto in situazioni diverse da quella statunitense, i problemi di costo, tempo, reperimento di personale atto ad impostare ed effettuare rilevazioni basate su outcome clinici sono elevatissimi anche qualora tutti i problemi definitivi sono superati in modo soddisfacente.

Una prima alternativa agli outcome clinici è data dall'utilizzo di outcome inerenti la qualità della vita che descrivono la funzionalità fisica o psichica

delle strutture) e risultato (indicatori di mortalità per patologia e trattamenti). Gli indicatori per area sono indicatori di processo inerenti l'utilizzazione delle strutture sanitarie e le ospedalizzazioni evitabili.

b) indicatori della *Joint Commission of Health Organization* (JCAHO) contenuti nella *National Library of Healthcare Indicators* (1997). Sono 225 indicatori di processo, clinici, di soddisfazione dell'utente, di qualità della vita, amministrativo-finanziari. Per ogni indicatore sono elencate caratteristiche statistiche e sanitarie, campi di applicazione, modo di costruzione, enti che li propongono e li verificano.

²⁵ Nella stragrande maggioranza gli indicatori sono frequenze relative di successi nelle cure. Per costruirli si tiene conto della gravità dei pazienti prima della cura "pesando" la frequenza dei singoli pazienti mediante le modalità ordinali dell'APR-DRG, riclassificazione dei DRG sulla base della gravità (vedi nota 54). Tuttavia per come sono costruiti gli indicatori non possono essere utilizzati né come variabili esplicative né come variabili dipendenti osservate a livello del singolo paziente ma solo come variabili esplicative inerenti le strutture ospedaliere. Inoltre essendo presentati a livello aggregato di ospedale non si può tener conto di eventuali loro correlazioni. Infine la menzionata procedura attraverso cui si ponderano le frequenze mediante gli indicatori di gravità dei pazienti desunto dall'APR-DRG non è raffinato sul piano statistico.

²⁶ I finanziatori del sistema quali le Assicurazioni, le *Health Maintenance Organizations*, lo stesso Governo attraverso *Medicare e Medicaid*, richiedono alle strutture erogatrici di servizi ospedalieri la consegna periodica di dati inerenti indicatori di qualità che consentono di effettuare benchmarking interni. Ad esempio, le compagnie di assicurazioni rivedono sia i rimborsi per gli ospedali sia i premi assicurativi dei pazienti in funzione della qualità rilevata a fronte del sistema di indicatori. Tuttavia i dati da loro raccolti sono indisponibili al pubblico in quanto connessi al profitto aziendale.

²⁷ Nel mese di febbraio del 1997 l'JCAHO ha annunciato l'iniziativa ORYX, programma di lungo periodo, avente lo scopo di inserire in futuro la rilevazione di indicatori e *outcome* all'interno del processo di accreditamento. L'intendimento è di validare sistemi di indicatori e outcome proposti dai singoli ospedali piuttosto che imporne uno centrale. A questo scopo si propone che ogni insieme di indicatori proposti da una struttura sanitaria sia esaminata dalla *Joint Commission's Advisory Council on performance Measurements* rispetto a criteri quali: tipologia (di processo e di risultato), caratteristiche banche dati, qualità dei dati, metodi di trasmissione dei dati. Per aiutare in questo percorso l'JCAHO ha cominciato a proporre periodicamente indicatori considerati cruciali (core measurement). Sono stati così indicati negli ultimi anni 40 indicatori rispetto a 5 problemi di salute fondamentali quali infarto, scompenso cardiaco, polmonite, complicanze chirurgiche, gravidanze.

generale o quella inerente ad aspetti particolari della vita del paziente²⁸. Tali *outcome* sono rilevabili in modo più semplice degli *outcome* clinici su scala ordinale attraverso opportuni questionari; tuttavia ognuno di loro è specifico per particolari patologie e la loro rilevazione comporta problemi simili a quelli descritti per gli *outcome* clinici.

Una terza alternativa è quella di rilevare *outcome* avversi o “eventi sentinella”²⁹. Tali *outcome* mettono in evidenza non solo le carenze professionali dei singoli ma anche le carenze strutturali e organizzative degli ospedali nel loro complesso e sono molto più facilmente rilevabili rispetto agli *outcome* clinici e di qualità della vita da dati di origine amministrativa.

Infine vi sono numerosi *outcome* che non rilevano né le condizioni cliniche dei pazienti, né la loro funzionalità, né eventi sentinella ma solo condizioni necessarie perché si verifichino conseguenze negative per i pazienti. Tali *outcome*, che potremmo denominare in generale come “*outcome* di struttura” sono senz’altro più imprecisi in quanto la loro esistenza non implica automaticamente situazioni di inefficacia ma sono anche più facilmente rilevabili in modo sistematico da dati amministrativi anche per misurare lo stato di salute del paziente prima della cura.

In conclusione l’estrema difficoltà a rilevare, prima e dopo l’erogazione del servizio, gli *outcome* clinici e di qualità della vita, per una pluralità di patologie, suggerisce il loro utilizzo solo nell’ambito di studi di nicchia inerenti particolari patologie quando si abbiano a disposizione dati di origine clinico epidemiologico provenienti da campioni costruiti secondo disegni sperimentali o quasi sperimentali. In casi più generali in cui si abbiano solo dati osservazionali provenienti da campioni non costruiti secondo disegni sperimentali o quasi sperimentali, oppure si voglia costruire un sistema di valutazione generale degli ospedali basato su un vasto complesso di patologie sulla base di dati amministrativi inerenti campioni di elevate dimensioni o la popolazione nel suo

²⁸ Ad esempio la *Functional Independence Measurement* è specifica per le patologie che indicano la riabilitazione (Tesio ed al.,1997), l’PSF-12 (Ware, 1993) si usa preferibilmente per malattie cronico-degenerative.

²⁹ Eventi insospettiti che comportino per il paziente la morte o danni seri fisici o psicologici (quali lesioni gravi e perdita o limitazione di funzioni), o rischi consistenti di morte e danni (atti non intenzionali, omissioni, atti che non raggiungono risultati attesi, variazioni nel processo assistenziale che non hanno conseguenza sugli outcome ma il cui verificarsi porta a una significativa possibilità di outcome avversi), condizioni di pericolo (circostanze specifiche che incrementano specificatamente la probabilità di outcome avverso in relazione alla patologia di cui è affetto il paziente) (Jcaho 2000) Il manuale di accreditamento dell’JCHAO contiene suggerimenti per rilevare eventi sentinella e di struttura e il modo per rilevarli; un programma specifico dell’JCHAO “*Sentinel Alert*” fornisce indicazioni per evitarli; il manuale di indicatori del 1997 già citato ne contiene un elenco.

complesso, è preferibile optare per outcome “sentinella” e di struttura rilevati prima e dopo l'erogazione del servizio.

2. LE FONTI INFORMATIVE, GLI OUTCOME, LE VARIABILI ESPLICATIVE: LE SDO³⁰

Quanto finora detto vale in termini generali. V'è ora da chiedersi come e' possibile costruire un sistema di valutazione basato su dati amministrativi per le strutture ospedaliere in Italia.

Nel nostro paese mancano rilevazioni sistematiche inerenti outcome clinici e sulle qualità della vita valide per tutto il territorio nazionale anche se vi sono interessanti raccolte a livello regionale³¹.

L'unica fonte informativa sistematica esistente in Italia a riguardo dei ricoveri ospedalieri è la scheda di dimissione ospedaliera (d'ora in poi SDO)³² Dalle SDO

³⁰ TABELLA 2
Scheda di dimissione Ospedaliera

1\2	Codice subcodice struttura	26	Data dimissione o morte
3\4	Anno di ricovero + n°pratica		
5	Mese e giorno di ricovero (mmgg)	21/23/25	Reparto trasferimento (1-2-3)
6	Regime di ricovero	27	Diagnosi principale alla dimissione (per DRG)
7	Reparto di ricovero	28/32	Altra condizione patologica (1-5)
8	Sesso	33	Data intervento chirurgico principale
9	Codice individuale	34	Intervento principale o parto
10	Luogo di nascita	35/37/39/41/43	Data altro intervento(1- 2- 3 -4 -5)
11	Data nascita (età)	36/38/40/42/44	Altro intervento (1-2-3-4-5)
12	Residenza anagrafica	45	Modalità di dimissione
13	Cittadinanza	46	Riscontro autoptico
14	Stato civile	47	Day hospital – giornate di presenza
15	Posizione nella professione	48	Day hospital – motivo di ricovero
16	Provenienza paziente	49	Peso alla nascita
17	Tipo di ricovero	50	Codice individuale della madre X
18	Onere della degenza	55	Tipo tariffa (“T” per tariffe non regionali)
19	Modalità di trauma	56	Importo totale prestazione (solo tariffe non regionali)
20/22	Dati trasferimento reparto (1-2-3)	57/58	Giornate non a carico del SSN
24			Destinazione del record

³¹ Ad esempio in Friuli sono disponibili banche dati che forniscono informazioni a livello del singolo paziente a riguardo di anagrafe assistiti, nascite, allattamento al seno, vaccinazioni, prescrizioni farmaceutiche, referti di anatomia patologica, prestazioni ambulatoriali,registri tumori,schede di morte, appropriatezza del ricovero,infezioni chirurgiche postoperatorie, lesioni da decubito. Tali banche dati non sono però collegate in una unico sistema informativo.

si possono ricavare informazioni su alcuni eventi sentinella, *outcome* di struttura, e su alcuni aspetti della gravità del paziente prima della cura. Nella tabella 3, 4, 5 sono elencati rispettivamente i più importanti *outcome*, variabili esplicative inerenti i pazienti, variabili esplicative inerenti gli ospedali ricavabili dalle SDO.

TABELLA 3

Outcome desumibili dalle SDO. Fonti: SDO, registri sale operatorie, cartelle cliniche

		TIPO OUTCOME	SCALA
<i>y1</i>	Tasso di mortalità intraospedaliera ³³	Clinico	Dicotomica
<i>y1a</i>	Tasso di mortalità intraospedaliera (morte evitabile)	Evento sentinella	Dicotomica
<i>y2</i>	Autodimissioni ³⁴	Evento sentinella	Dicotomica
<i>y3</i>	Trasferimenti tra strutture ³⁵	Struttura	Dicotomica
<i>y4</i>	Ricoveri ripetuti per lo stesso MDC ³⁶	Struttura	Dicotomica
<i>y4a</i>	Ricoveri ripetuti per infezioni chirurgiche ³⁷	Evento sentinella	Dicotomica
<i>y5</i>	Ritorno in sala operatoria ³⁸	Struttura	Dicotomica
<i>y6</i>	Appropriatezza ricovero ³⁹	Struttura	Continua
<i>y7</i>	Degenza preoperatoria ⁴⁰	Struttura	Continua
<i>y8</i>	Ritorno in terapia intensiva ⁴¹	Struttura	Dicotomica
<i>y9</i>	Complicanze del ricovero ⁴²	Struttura	Dicotomica
<i>y9a</i>	Complicanze post-operatorie ⁴³	Evento sentinella	Dicotomica
<i>y10</i>	Ritardo nel trattamento ⁴⁴	Evento sentinella	Continua
<i>y11</i>	Complessità casistica ⁴⁵	Struttura	Continua
<i>y12</i>	Trasformazione del regime di ricovero da <i>day hospital</i> ad ordinario	Struttura	Dicotomica

³² Le informazioni riguardanti *outcome* ed indicatori descritti in questo paragrafo sono tratte da JCAHO (2000), dal Ministero della Salute, dalla Regione Lombardia.

³³ Per avere dati inerenti la mortalità entro 30 giorni dalle dimissioni occorre integrare i dati delle SDO con quelli ISTAT sulla mortalità. Quando non sono eventi sentinella e quindi decessi evitabili (nota 34) si hanno due possibilità. Si hanno decessi che provengono da condizioni in prevalenza di tipo chirurgico (caratterizzati da un rischio relativamente elevato di mortalità intraospedaliera) e da condizioni in prevalenza di tipo medico (in cui è attesa una bassa mortalità).

³⁴ Sono eventi sentinella se avvengono contro il parere dei sanitari, da distinguere da quelle in condizioni terminali, da assimilare al fenomeno della mortalità intraospedaliera.

³⁵ I trasferimenti che avvengono oltre una determinata soglia di giornate di degenza, con maggiore probabilità sono associabili a complicanze del ricovero o ad un processo assistenziale non ottimale. Invece in alcuni casi, i trasferimenti tempestivi sono indice di un buon processo assistenziale.

³⁶ Vengono intesi i ricoveri ripetuti dello stesso paziente nella medesima *Major Diagnostic Category* (MDC) (vedi nota 54). Quando siano sintomo dell'insorgere di patologie adiacenti a quella originaria, sono pertanto potenzialmente riconducibili a un non corretto inquadramento del paziente e quindi ad una non efficacia della cura.

TABELLA 4

Variabili esplicative pazienti

		Scala
x_1	Sesso ⁴⁶	Dicotomica
x_2	Età paziente ⁴⁷	Discreta
x_3	Tipo di ricovero ⁴⁸	Dicotomica
x_4	Stato civile	Ordinale
x_5	Professione ⁴⁹	Ordinale
x_6	Indicatore di urgenza ⁵⁰	Dicotomica
x_7	Provenienza	Sconnessa
x_8	Durata di degenza ⁵¹	Discreta
x_9	Scostamento dalla degenza media ⁵²	Continua
x_{10}	Degenza pre-intervento	Discreta
x_{11}	Gravità paziente ⁵³	Ordinale

³⁷ Sono eventi sentinella relativi a pazienti riammessi nello stesso ospedale con una diagnosi principale sul sito clinico già operato.

³⁸ Si tratta di pazienti che rientrano in sala operatoria per un nuovo intervento che può avvenire sullo stesso sito, nel caso di una complicanza in sito (es.: deiscenza della ferita chirurgica, "malpractice", ecc.), o comunque per motivi derivanti dall'intervento iniziale.

³⁹ Attraverso il D.P.C.M. 29 novembre 2001 "Definizione dei livelli essenziali di assistenza" Gazzetta Ufficiale N.33 dell'8.2.2002- Supplemento Ordinario N.26, il Ministero ha definito "inappropriati" i casi trattati in regime di ricovero ordinario o in day hospital che gli ospedali possono trattare con identico beneficio per il paziente e con minore impiego di risorse e quindi ha individuato un elenco di 43 DRG "ad alto rischio di inappropriatazza". I DRG sono raggruppamenti di diagnosi secondo il principio di consumo di risorse comunemente usati nel nostro paese dal 1995. Lo scostamento dalla mediana dei DRG, rilevato a livello di ospedale, è perciò un indicatore di appropriatezza di tipo continuo.

⁴⁰ La tendenza generale è di ridurre sempre di più la degenza preoperatoria dei ricoveri ordinari mediante un prericovero programmato per l'esecuzione degli esami preoperatori di routine, per aumentare la produttività.

⁴¹ I ritorni in terapia intensiva rappresentano una *proxy* delle complicanze e/o di un peggioramento delle condizioni cliniche avvenute durante il trattamento del paziente.

⁴² Per separare le complicanze dalle comorbidity, mettendo in luce soprattutto le complicanze correlabili a errori medici occorre confrontare la SDO con il contenuto delle cartelle cliniche, opportunamente campionate utilizzando all'uopo un apposito algoritmo di cattura automatica delle complicanze denominato CSP – *complication screening program*.

⁴³ E' inerente pazienti chirurgici che sviluppano complicanze del sistema nervoso centrale o periferico e infezioni nosocomiali postoperatorie sul sito chirurgico, come seconda diagnosi dopo operazioni chirurgiche.

⁴⁴ Diviene continua se si calcola il ritardo in termini di tempo d'attesa.

⁴⁵ La complessità della casistica è una *proxy* dell'impegno richiesto dal caso che viene costruita utilizzando la classificazione DRG valutata in termini di risorse impiegate per la cura. Gli indicatori dell'HCUP degli Usa ipotizzano che la qualità di un ospedale sia direttamente proporzionale alla produzione di casistica a peso elevato. Il case mix medio è dato dal numero di pazienti affetti da

TABELLA 5

*Variabili esplicative struttura*⁵⁴

		SCALE
Z ₁	Dimensione ospedale:posti letto	Discreta
Z ₂	Numero casi per ospedale	Discreta
Z ₃	Tipologia: % posti letto per AREA	Continua
Z ₄	Tipologia:% costi per AREA	Continua
Z ₅	%costi personale sul totale costi per OSP/AREA	Continua
Z ₆	Costo per posto letto utilizzato per OSP/AREA	Continua
Z ₇	% utilizzo posti letto accreditati per OSP/AREA	Continua
Z ₈	N [^] medio ore utilizzo sala operatoria	Discreta
Z ₉	% pazienti operati su dimessi	Continua
Z ₁₀	Degenza media per OSP	Discreta
Z ₁₁	N [^] orario pazienti in pronto soccorso ⁵⁵	Discreta
Z ₁₂	Case mix medio	Continua

una patologia per l'indice quantitativo ad esso legato fornito dal DRG e indica il grado di specializzazione dell'ospedale e la capacità di curare patologie complesse.

⁴⁶ Interessa tutti gli *outcome*. Numerose prestazioni chirurgiche sono completamente rivolte al sesso femminile (come quelle di ostetricia e ginecologia) o al sesso maschile (es.: tumori maligni e benigni della prostata, ecc.).

⁴⁷ Interessa tutti gli *outcome*. E' l'età del soggetto in corrispondenza dell'evento sanitario di interesse.

⁴⁸ Permette di distinguere se il paziente è stato ricoverato in degenza ordinaria o in *day hospital*.

⁴⁹ E' una variabile non rilevante di per sé, ma in quanto permette di costruire altre variabili: in particolare, in congiunzione con la scolarità, permetterebbe di identificare lo stato socio-economico del soggetto.

⁵⁰ E' un indicatore che, per l'area chirurgica, permette di individuare la gravità del paziente all'ingresso.

⁵¹ Poiché la durata del ricovero influenza il pagamento, ed in particolare definisce il prezzo pieno del ricovero, nel caso dei ricoveri brevi può essere oggetto di fenomeni di opportunismo.

⁵² In letteratura la "*Average Length of Stay*" viene considerata una *proxy* della complessità del caso che, in un sistema tende alla contrazione per migliorare l'efficienza.

⁵³ Esistono diversi modi per classificare le diagnosi in funzione della gravità partendo dai DRG. Le principali sono : MDC "*Major diagnostic category*" raggruppa in 25 sottogruppi i DRG secondo omogeneità di apparato, APR-DRG utilizza sia il principio del consumo di risorse sia il principio della gravità ed è usato nel nostro paese a scopo sperimentale e richiede adozione di uno specifico software; Disease Staging: è un raggruppamento di diagnosi basato sulla gravità delle malattie ed è usato nel nostro paese solo a scopi sperimentali e richiede anch'esso un apposito software.

⁵⁴ Gli indicatori Z₁-Z₆ rappresentano gli input, o i fattori produttivi delle strutture sanitarie; gli indicatori Z₇-Z₁₂ sono gli output inerenti l'ospedale che vengono prodotti in funzione dell'uso di fattori produttivi. Tanto sono maggiori le quote dei posti letto accreditati, il numero medio di ore di utilizzazione delle sale operatorie, la quota di operati su dimessi e tanto minori sono i costi per posto letto e la degenza media, quanto maggiore è l'efficienza.

⁵⁵ Il numero di accessi in pronto soccorso è indice di carico non programmabile per l'ospedale

2.1 Un esempio per il caso lombardo analisi dell'efficacia relativa all'area chirurgica di ospedali

Per verificare la metodologia proposta si effettua una analisi dell'efficacia relativa degli ospedali lombardi. Si analizzano gli outcome $y_1, y_2, y_3, y_4, y_5, y_6, y_7, y_{11}$; in funzione delle variabili esplicative inerenti pazienti e strutture elencate nelle tabelle 4 e 5⁵⁶. La fonte statistica è data dalle SDO dell'anno 2001 inerenti 650.000 ricoverati negli ospedali lombardi nell'area chirurgica integrate per ciò che concerne le variabili riguardanti gli ospedali da dati amministrativi a disposizione della Regione Lombardia. Per i primi cinque outcome, definiti a livello di paziente e ospedale si utilizza il modello multilevel con interazioni secondo le formule (1) e (9), di tipo logistico, con distribuzione degli outcome dicotomica, distribuzione degli errori extrabinomiale, distribuzione delle variabili casuali u_j normale. Le equazioni relative a y_6, y_7, y_{11} , outcome definiti a livello di ospedali, sono analisi delle varianze. I risultati, espressi in termini di odds ratio sono stati ottenuti mediante il metodo della "correlation-based feature selection" (CFS), (Hall 1998)⁵⁷. I risultati più importanti sono contenuti nella tabella 6. La variabilità di y_1 , è maggiore entro gli ospedali che fra gli ospedali a significare l'assenza di fenomeni di gravi inefficacie nella regione (gli odd ratios sono tutti compresi tra 0.98 e 1.02)⁵⁸. Tra le variabili esplicative spiccano x_2 (età), x_6 (urgenza), x_8 (durata degenza del paziente), z_{10} (degenza media per ospedale) indicatori di gravità all'ingresso dell'ospedale, a segnalare che variabili esplicative riescono a far sì che la mortalità espressa dalla distribuzione degli errori u_j , su cui effettuare i confronti di efficacia relativa sia realmente risk adjusted, nel senso

⁵⁶ Alcune variabili esplicative sono riclassificate in termini dicotomici nel seguente modo: x_7 provenienza (Lombardia/ Extra Lombardia) x_8 Durata degenza (≤ 10 giorni; > 10 giorni); x_9 spostamento dalla degenza media (> 1 giorno; ≤ 1 giorno) x_{10} Degenza preintervento (> 1 giorno; ≤ 1 giorno). Altre sono riclassificate in modo ordinale x_7 Dimissione ospedale (≤ 200 ; 201- 500, ≥ 500); z_{14} N° orario pazienti in pronto soccorso (0-2; 3-4; 5-6; 7-8; > 8). Le variabili z_5 z_6 z_7 sono rilevate sia per gli ospedali nel loro complesso che per le aree chirurgiche (z_{13} ; z_{14} ; z_{15})

⁵⁷ Il CFS, metodo di "machine learning" usa un algoritmo di ricerca basato sulla misura della bontà di sottoinsiemi di variabili come predittori. Il criterio è il seguente $G_s = \frac{k\bar{r}_{ci}}{\sqrt{k + k(k-1)\bar{r}_i}}$

Dove: k è il n° di variabili X_k il sottoinsieme di variabili considerate, r_{ci} è la correlazione media delle X_k con l'outcome y_j , r_{ii} è l'intercorrelazione media tra le X_k . La correlazione è definita come segue:

$$C(y/X) = \frac{H(y) - H(y/X)}{H(y)} \text{ dove } H(y) = -\sum p(y) \log(p(y)) \text{ ed } H(y/X) = -\sum_x p(x) \sum p(y|x) \log(p(y|x))$$

La procedura esegue un'iterazione sui gruppi di variabili e deriva il sottoinsieme X di y più predittivo.

⁵⁸ La mortalità non considera i morti fuori dall'ospedale entro 30 giorni dalla dimissione non disponibile.

indicato nel paragrafo 2. Ciò vale anche, in diverso modo per gli altri outcome y_2, y_3, y_4, y_5 su cui si effettua l'analisi di efficacia relativa tra gli ospedali data l'importanza di x_6 (urgenza) ed x_8 (durata degenza del paziente) tra le loro variabili esplicative. La variabilità dell'outcome y_2 (dimissioni volontarie) tra ospedali è bassissima; tra le sue variabili esplicative spicca x_8 (durata della degenza) a mostrare situazioni di disagio provenienti da ricoveri prolungati. Il valore del coefficiente relativo all'interazione x_8-z_{13} (costi del personale per area chirurgica) suggerisce che i disagi possano essere dovuti anche a una gestione non oculata delle risorse umane nell'area chirurgica che determina una ricaduta negativa sull'utenza. La forte variabilità dei trasferimenti fra strutture (y_{13}) indica che ci si trasferisce verso ospedali più capaci di cure a fronte di patologie complesse e gravi per motivi di urgenza e gravità, come mostra il fatto che la variabile esplicativa x_6 è la più importante. Dal fatto che i trasferimenti sembrano avvenire principalmente verso strutture in cui la degenza pre-intervento è maggiore di un giorno (x_{10}), si conclude che si tratta di trasferimenti da piccole strutture non specializzate (es.: i vecchi ospedali di zona, piccole case di cura, ecc.) verso strutture più ottimali (grandi ospedali/ospedali specializzati). Anche la variabilità di y_4 (ricoveri ripetuti) è elevata: il fatto che tra le variabili esplicative più importanti vi sia x_8 (durata degenza del paziente) mostra che tale comportamento nasce quando si ha difficoltà ad uscire dalla situazione acuta sotto il profilo patologico già al primo ricovero; l'importanza dell'interazione di x_6 con z_8 mostra che tale fenomeno riguarda soprattutto gli ospedali specializzati nell'area chirurgica. Per ciò che concerne y_5 (ritorno in sala operatoria), bassa è invece la variabilità tra ospedali. Dato il valore dell'*odd's ratio* relativo a z_{13} (% costi personale in area chirurgica), di gran lunga inferiore ad 1, l'equazione mostra come il ritorno in sala operatoria sembri legato oltre che alla gravità, a problemi d'efficienza dell'ospedale. Le distorsioni nell'appropriatezza dei ricoveri y_6 sembrano direttamente correlate al crescere dell'attività dell'ospedale, come dimostra il fatto che le variabili esplicative più importanti sono z_5 (tasso di occupazione dei letti accreditati) e z_{10} (degenza media); La degenza media preoperatoria (y_7) aumenta per ospedali poco specializzati nell'area chirurgica (z_3), il che riporta l'attenzione ancora sui piccoli ospedali meno attrezzati. Per ciò che concerne la complessità delle casistiche (y_8) si osserva che tra le variabili esplicative determinanti risultano z_5 (% di costi dell'area chirurgica sul totale) e z_7 (% di utilizzo di posti letto accreditati) a significare che la complessità della casistica aumenta in ospedali maggiormente specializzati e quindi maggiormente accreditati e con reparti chirurgici più importanti.

TABELLA 6
Coefficienti più importanti delle variabili esplicative nei modelli multilevel

	Coefficiente di correlazione intraclassi	Valore atteso	Intervallo di confidenza $\hat{\beta}_j$	p-value
y_1	0.2915			
x_2		2.205	2.18 – 2.23	<0.0001
x_6		0.0619	0.04 – 0.09	<0.0001
x_8		1.097	1.086 - 1.10	<0.0001
z_{13}		1.405	1.208 - 1.633	<0.0001
$x_6^*z_{13}$		1.01	1.036 - 1.166	0.0017
$x_8^*z_{13}$		0.992	0.99 – 0.993	<0.0001
y_2	0.057			
x_8		0.007	0.006-0.008	<.0001
$x_8^*z_{13}$		2.083	1.547-2.806	<.0001
y_3	0.617			
x_6		0.16	0.145 - 0.176	<.0001
x_{10}		1.067	1.06 - 1.073	<.0001
y_4	0.563			
x_6		0.204	0.062-0.672	.01
x_8		4.011	3.775-4.261	<.0001
$x_6^*z_8$		3.157	0.908-10.984	0.0707
y_5	0.119			
x_2		1.006	1.004-1.008	<.0001
z_6		0.162	0.074-0.354	<.0001
$x_2^*z_{13}$		1.019	1.014-1.024	<.0001
$x_8^*z_{13}$		1.111	1.097-1.126	<.0001
y_6				
z_5		0.804	0.505-1.103	<.0001
z_{10}		7.44	4.286-10.593	<.0001
y_7				
z_3		-0.403	-0.674-0.132	0.0041
z_{10}		0.227	0.2021-0.253	<.0001
y_8				
z_4		0.3743	0.210 – 0.538	<.0001
z_7		0.2428	0.217 – 0.268	<.0001

A questo punto si effettuano i confronti fra ospedali per stabilire l'efficacia relativa degli ospedali per le prime cinque equazioni. I risultati dei confronti tra gli ospedali appaiono nelle figure 1,2,3,4,5. In questo caso essendo la rilevazione effettuata non su campioni ma sull'universo delle popolazione si possono effettuare confronti sui valori attesi dagli errori piuttosto che sugli estremi dei loro intervalli, per gli outcome y_1, y_3, y_4, y_5 . Gli ospedali più piccoli e meno specializzati si situano nella parte terminale delle distribuzioni degli *outcome*. Il contrario avviene per gli ospedali più grandi o più specializzati. Ciò sembra suggerire che la dimensione e la specializzazione dell'ospedale risulta determinante per assicurare condizioni di efficacia. L'*outcome* y_2 risulta invece

influenzato anche da altre componenti perché, come si può supporre anche in chiave teorica le dimissioni volontarie sono anche dovute a ragioni di efficienza.

Per avere conferme rispetto a queste conclusioni si effettua una verifica delle concordanze, delle graduatorie generate dai diversi *outcome* mediante il coefficiente di *Spearman* (tabella 7).

TABELLA 7
Analisi dei coefficienti di Spearman nell'area chirurgica

	Mortalità intraospedaliera	Autodimissioni	Trasferimenti tra strutture.	Ricoveri Ripetuti	Ritorni Sala Operatoria
Mortalità intraospedaliera	-	0.18	0.23	0.09	0.08
Autodimissioni		-	0.62	0.11	-0.05
Trasferimenti tra strutture			-	0.03	-0.04
Ricoveri ripetuti				-	0.55

La graduatoria basata sulla mortalità intraospedaliera non è significativamente correlata con le altre; la graduatoria basata sulle dimissioni volontarie ha una correlazione forte con quella basata sui trasferimenti tra strutture ($r=0.62$); quella basata sui ricoveri ripetuti ha una correlazione forte con la graduatoria basata sui ritorni in sala operatoria ($r=0.55$). Se ne deduce che l'efficacia degli ospedali può essere spiegata da tre componenti. La prima è quella individuata dalle dimissioni volontarie e trasferimenti tra ospedali per essere curato meglio il paziente decide di trasferirsi, ed è trasferito, in strutture più idonee ad assisterlo. I ricoveri ripetuti sono invece legati a ritorni in sala operatoria e mostrano una strategia di cura legata ad una determinata struttura. La mortalità sembra mettere in luce un aspetto originale dell'efficacia rispetto agli altri *outcome*.

La complessità del problema necessita ulteriori approfondimenti in particolare nella sintesi dei risultati delle analisi effettuate. A tale proposito sono in corso studi per sintetizzare gli indicatori utilizzati considerando le caratteristiche degli ospedali dal punto di vista dimensionale e di complessità della casistica afferente. Lo scopo finale è quello di realizzare una strategia di analisi che dia luogo ad un sistema valutativo che sia il più possibile "*coeteris paribus*"⁵⁹.

⁵⁹ Dato ad esempio da: utilizzo di un insieme di variabili esplicative capace di rappresentare in modo adeguato la variabilità degli ospedali; utilizzo dei residui per costruire graduatorie specifiche per indicatore; sintesi degli indicatori possibilmente stratificando per dimensione dell'area e pesando per complessità della casistica;

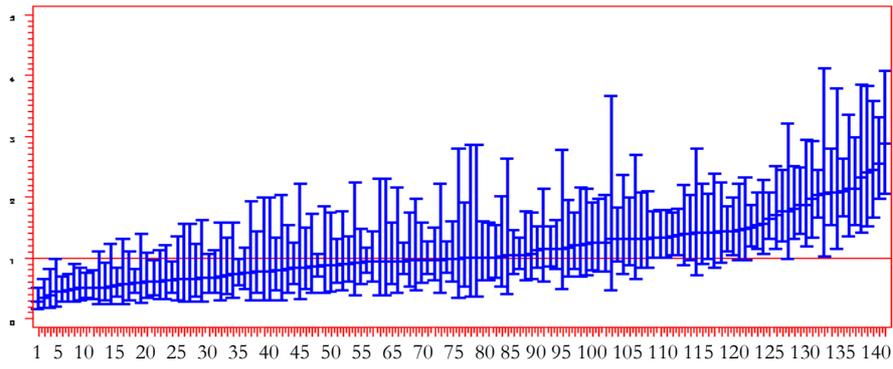


Figura 1 - y_1 mortalità intraospedaliera (intervalli di confidenza per ospedali)

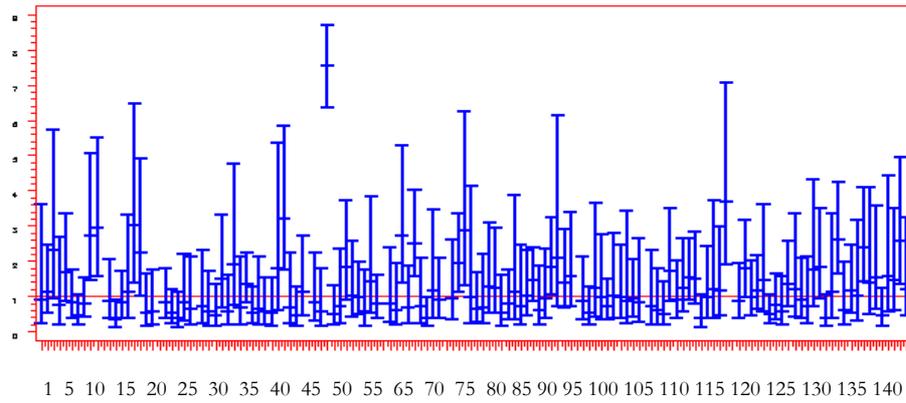


Figura 2 - y_2 dimissioni volontarie intervalli di confidenza per cosp

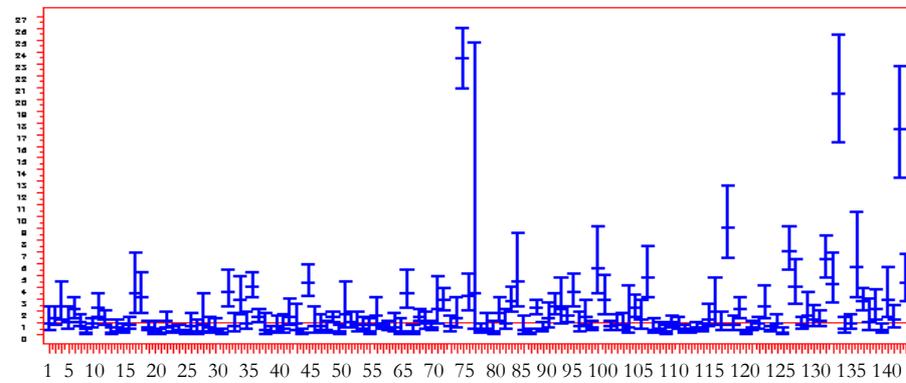


Figura 3 - y_3 trasferimenti tra strutture intervalli di confidenza per cosp

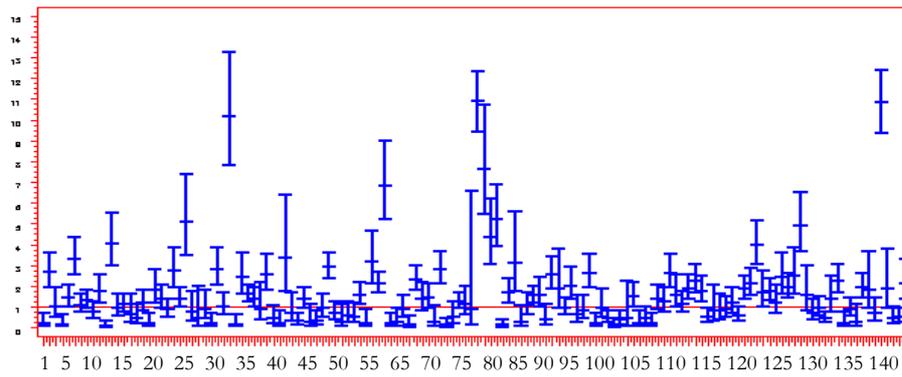


Figura 4- y_4 ricoveri ripetuti per lo stesso MDC

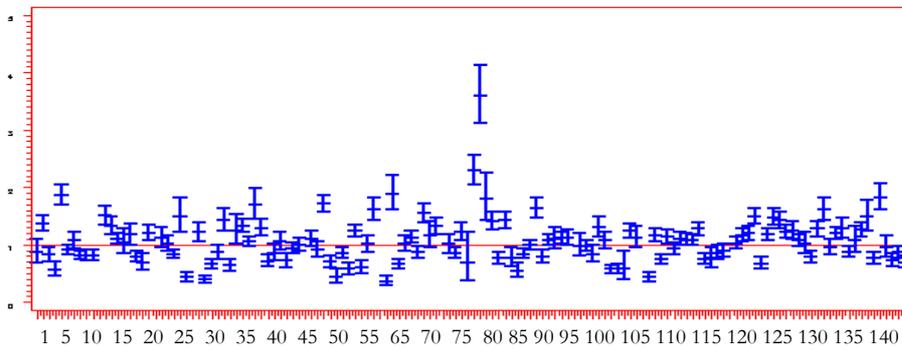


Figura 5- y_5 a ritorno in sala operatoria

Università di Milano Bicocca

Noustat S.R.L.

A. S. L. Brescia

Giorgio Vittadini

Maurizio Sanarico

Camillo Rossi

RIFERIMENTI BIBLIOGRAFICI

- D.J. AIGNER, C.A.K. LOVELL, P. SCHMIDT (1977), "Formulation and Estimation of Stochastic Frontier Production Function Models", *Journal of Econometrics*, vol. 6, pp. 21-37
- J. ALBERT, S. CHIB (1997), "Bayesian tests and model diagnostics in conditionally independent hierarchical models", *Journal of the American Statistical Association* 92, 916-25
- T.G. BOND, C.M. FOX (2002), "Applying the Rasch Model: Fundamental Measurement in the Human Sciences", Lawrence Erlbaum Associates Publishers, Mahwah, New Jersey
- C. DAGUM, (1987), "Measuring the economic Affluence between populations of income receivers", *Journal of business & economic statistics*, 5,1 pp 5-11
- J. MD DALEY, MD.JENCKS, D. DRAPER, G.LENHART, N. THOMAS, J. WALKER (1988), "Predicting Hospital -Associated Mortality for Medicare Patients, A method for patients with stroke, pneumonia, acute myocardial infarction, and congestive heart failure", *JAMA*, Dec 23/30, 1988 Vol 260, N. 24, 3617,3624.
- P. DAVIS, B. GRIBBEN (1995), "Rational prescribing and interpretational variation: a multilevel approach", *International Journal of Technology Assessment in Health Care* 11, 428-442
- L. D'AMBRA. (2001), "Metodi Statistici per la misurazione della Customer Satisfaction", *Atti della Scuola della Società Italiana di Statistica*, Roma
- A. DONABEDIAN (1990), "La qualità dell'assistenza sanitaria. Principi e metodologie di valutazione", La Nuova Italia Scientifica, Roma
- DUBOIS RW, BROOK RH, ROGERS WH (1988), "Adjusted hospital death rates: A potential screen for quality of medical care", *American Journal of Public Health* 77, 1162-7
- C. DUNCAN, K. JONES, G.. MOON (1998), "Context, composition and heterogeneity: using multilevel models in health research", *Social Science and Medicine* 46, 97-117
- A. EPSTEIN (1995), "Performance reports on quality-prototypes, problems and prospects", *New England Journal of Medicine* 333, 57-61
- A. EPSTEIN, FC. SCHNEIDER (1997), "Influence of cardiac-surgery performance reports on referral practices and access to care", *New England Journal of Medicine* 335, 251-6
- H. GOLDSTEIN, A.H. LEYLAND (2001), "Further developments" in Goldstein H., Leyland A.H. (eds.) *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 175-186
- H. GOLDSTEIN, D.J. SPIEGELHALTER (1996), "League tables and their limitations: statistical issues in comparison of institutional performance", *JRSS*, Ser.A,159,3, pp. 385-443
- H. GOLDSTEIN (2000), "Random effects meta-analysis of trials with binary outcomes using multilevel models" in MlwiN, *Multilevel Modelling Newsletter*, vol. 11, n. 1
- E. GORI (1992), "La valutazione dell'efficienza e dell'efficacia dell'istruzione", *Atti della XXXVI Rinnione Scientifica della SIS*, Vol. 1, pp. 219-229
- E. GORI, C. ROSSI, L. GRASSETTI (2001), "La valutazione dell'efficienza delle strutture sanitarie della regione Lombardia", in *Statistica applicata*, in corso di pubblicazione
- E. GORI., G. VITTADINI (1999), "La valutazione dell'efficienza e dell'efficacia dei servizi alla persona. Impostazione e metodi", in *Qualità e valutazione nei servizi di pubblica utilità*, Gori E, Vittadini G (a cura di), Etas Libri, Milano, pp. 121-241
- K. HAAGEN, G. VITTADINI (1998), "Restricted Regression Component Decomposition", *Metron*, vol. 56, pp. 53-75
- M.A. HALL (1998), "Correlation-based feature selection for Machine Learning, PhD thesis", Dept. of Computer Science, University of Waikato, Hamilton. New Zealand, 1998.
- M.J.R. HEALY (2001), "Multilevel Data and Their Analysis" in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 1-12
- D. HEDEKER, RD. GIBBONS (1994), "A random effects ordinal regression model for multilevel analysis", *Biometrics* 50, 933-44
- JJ. HOX (1995), "Applied Multilevel Analysis", TT-Publikaties, Amsterdam
- JCAHO (2000), Joint Commission on Accreditation of healthcare Organizations, www.jcaho.org

- S.F. JENCKS, J. DALEY, D. DRAPER, N. THOMS, G. LENHART, J. WALKER (1988), "Interpreting Hospital Mortality Data. The Role of Clinical Risk Adjustment", *Journal of the American Medical Association* 260, 3611-16
- CM. KUNIN, KS. JOHANSEN, AM. WORNING, FD, DASCHNER (1990), "Report of a symposium on use and abuse of antibiotics worldwide", *Reviews of Infectious Disease* 12, 12-19
- I. LANGFORD, R.J. DAY (2001), "Poisson Regression" in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp.45-58
- I. LANGFORD, T. LEWIS (1988), "Outliers in multilevel data", *Journal of the Royal Statistical Society A* 161, 121-60
- A.H. LEYLAND, F.A. BODDY (1998), "League tables and acute myocardial infarction", *Lancet* 351, 555-8
- T. LEWIS, I. LANGFORD (2001), "Outliers, Robustness and the Detection of Discrepant Data" in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp.75-92
- X.S. LI (1997), "Multilevel Modelling for Discrete Responses" in *Epidemiological Field Study*, PhD Dissertation, School of Public Health, West China University of Medical Sciences
- EC. MARSHALL, D.J. SPIEGELHALTER (1999), "League tables of in vitro fertilisation clinics: How confident can we be about the rankings?", *British Medical Journal*, 317, 1701-04
- EC. Marshall, D.J. Spiegelhalter (2001), "Institutional Performance", in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 127-142
- C. MC CULLOCH, A.H. LEYLAND, F.A. BODDY (1997), "Understanding the relationship between length of stay and readmission rates for hospitals (progress report for the king's fund comparative Database Initiative)" University of Glasgow, Public Health Research Unit.
- A. MCLEOD (2001, "Multivariate Multilevel Models", in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 59-74
- CN. MORRIS, CL. CHRISTIANSEN (1996), "Hierarchical models for ranking and for identifying extremes, with applications", in Bernardo JO, Berger JO, Dawid, AP, AFM (eds) *Bayesian Statistics 5*, Oxford University Press, pp.277-97
- O. O'DONNELL, C. PROPPER, R UPWARD (1993), "United Kingdom", In: Van Doorslaer E, Wagstaff A, Rutten F (eds) *Equity in the Finance and Delivery of Health Care: An International Perspective* Oxford University Press, pp.237-261
- A. PAGANO, C. ROSSI (1999), "La valutazione dei servizi sanitari", in Gori E., Vittadini G. Eds. *Qualità e Valutazione nei servizi di pubblica utilità*, Serie gestione d'impresa-direzione, Etas, Milano
- A. PARASURAMAN, V. ZEITHMAL, L. BERRY (1985, "A Conceptual Model of Service Quality and its Implication for Future Research", *Journal of Marketing*, 49.
- J. RASBASH, W. BROWNE (2001), "Modelling Non-Hierarchical Structures" in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 93-106
- N. RICE (2001), "Binomial Regression" in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 27-44
- N. RICE, A. LEYLAND (1996), "Multilevel models: applications to health data", *Journal of Health Services Research and Policy* 1, pp. 154-64
- S. RICHARDSON, P.J. GREEN (1997), "On Bayesian analysis of mixtures with an unknown number of components", *Journal of the Royal Statistical Society B* 59, pp. 731-92
- L. RIZZI, C. FRANCESCUTTI, G. SIMON, G. BORGNOLO (1999), "Valutazione della qualità e sistemi di incentivazione nei servizi sanitari: il caso della regione Friuli Venezia Giulia", in Gori E., Vittadini G., *Qualità e valutazione nei servizi di pubblica utilità*, pp.295-310
- A. SCOTT, A. SHIELL (1997), "Do fee descriptors influence treatment choices in general practice? A multilevel discrete choice model", *Journal of Health Economics* 16, pp. 323-42
- M. SMANS, CS. MUIR, P. BOYLE (1992), "Atlas of Cancer Mortality in the European Economic Community", International Agency for Research on Cancer, Lyon

- L. TESIO, L. PERUCCA, M.A. BATTAGLIA, F. FRANCHIGNONI (1997), "Quality assessment of the FIM (Functional Independence Measure) Ratings Through Rasch Analysis", *Europa medicophysica*, 33,2, pp. 55-57
- N. THOMAS, N.T. LONGFORD, J.E. ROLPH (1994), "Empirical Bayes methods for estimating hospital specific mortality rates", *Statistics in Medicine* 13, pp. 889-903
- G. VITTADINI (1998), "On the validity of the indeterminate latent variables in the Lisrel Model", *Communications in Statistics, Theory and Methods*, vol. 17, n.3
- G. VITTADINI (1999), "Analysis of qualitative variables in structural models with unique solutions", in *Classification and Data Analysis- Theory and Application*, Vichi M, Opitz O (eds), Springer-Verlag, pp. 203-210
- G. VITTADINI (2000), "On the use of Multivariate Regression models in the context of Multilevel Analysis", in *Advances in Classification and Data Analysis*, Borra S, Rocci R, Vichi M, Schader M. (eds), Springer-Verlag, pp.225-232
- G. VITTADINI, M. SANARICO (2003), "Metodologia per lo studio dell'efficacia relativa in sanità", in *Statistica Applicata*, in corso di pubblicazione
- M. YANG (2001), "Multinomial Regression" in Goldstein H., Leyland A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 107-126
- A. WAGSTAFF, E. VAN DOORSLAER (1993), "Equity in the finance and delivery of health care: concepts and definitions", in: *Equity in the Finance and Delivery of Health Care: an international perspective*, Van Doorslaer E., Wagstaff A, Rutten F. (eds). Oxford University Press., 7-19
- J.E. WARE (1993), "Sf-36, Health Survey. Manual and interpretation". *Guide*, The Health Institute, New England Medical Center
- H. WOLD (1982), "Soft Modeling: The Basic Design and Some Extensions", in *System under indirect observation: causality, structure, prediction*, K.G. Jöreskog and H. Wold, Eds. V. 2, pp.1-54
- WRIGHT, MASTERS (1982), "Rating Scale Analysis", Mesa Press, Chicago
- A. ZANELLA (2000), "Valutazioni e modelli interpretativi di customer satisfaction: ricerca di un quadro complessivo", *Convegno intermedio, Processi e Metodi Statistici di valutazione*, SIS 2001, pp. 113-120

RIASSUNTO

Recenti sviluppi nella metodologia statistica per la valutazione dell'efficacia degli ospedali

In questo lavoro si presentano i recenti sviluppi nella metodologia statistica per la valutazione dell'efficacia degli ospedali. Il modello statistico usualmente proposto a questo scopo è il modello *multilevel*. Perciò innanzitutto si presentano le proposte di tipo metodologico inerente tale modello applicato al problema dell'efficacia quali la natura e la distribuzione degli indicatori osservati e degli *outcome*, le variabili causali inerenti l'efficacia degli ospedali, il modello nel suo complesso. Secondariamente si traggono conclusioni a riguardo del dibattito sulla scelta degli *outcome* e degli indicatori osservati capaci di misurare efficacia e condizione dell'utente prima della cura.

Uno studio inerente i pazienti dell'area chirurgica negli ospedali lombardi nel 2000 illustra i vantaggi operativi dell'approccio proposto.

SUMMARY

On the recent developments in statistical methodology concerning the efficacy of hospitals

In this paper the recent developments in methodology concerning the efficacy of hospitals are presented. The statistical model used in order to examine efficacy is the multilevel model.

First of all, we present new proposals regarding the nature and distribution of the observed indicators and of the outcomes, the causal variables concerning hospital efficacy, and the model as a whole. Secondly, we examine the debate on the choice of outcomes and observed indicators able to measure respectively efficacy and patient conditions preceding admission to hospital.

A study concerning surgery patients in Lombard hospitals in 2000 illustrates an example of the operative and innovative consequences of our approach.