

VALUTAZIONE DELL'EFFICACIA DELLE STRUTTURE SANITARIE

Introduzione

Giorgio Vittadini

Quando le prestazioni delle strutture sanitarie sono valutate senza considerare il case mix delle caratteristiche dei pazienti non solo si possono dare giudizi scorretti ma si può anche essere indotti a praticare una selezione avversa verso i pazienti più gravi al fine di non peggiorare la valutazione.

Per superare tali problemi si è proposto di analizzare l'efficacia relativa delle strutture sanitarie dopo aver attenuato l'effetto delle variabili di case mix mediante modelli logistici regressivi o multilevel. Attraverso tali modelli di "risk adjustment" il confronto fra diverse strutture sanitarie viene effettuato "ceteris paribus" sulla base di outcome "corretti" rispetto agli effetti del case mix. Delle direzioni in cui si stanno raffinando e generalizzando le metodologie statistiche per la valutazione dell'efficacia delle strutture sanitarie si occupano Vittadini; Carabalona, Rossi nel loro lavoro. In quest'ottica un primo sviluppo riguarda aspetti metodologici riguardanti gli indicatori che danno origine agli outcome sanitari, e le ipotesi avanzate sulle loro distribuzioni, le particolarità statistiche dei modelli utilizzati. In particolare assume la crescente importanza la scelta delle metodologie secondo cui trasformare in punteggi quantitativi le modalità di un insieme di indicatori originati da test e questionari con risposte qualitative.

Una vastissima discussione esiste anche a riguardo del tipo di outcome sanitari e variabili di case mix da utilizzare per valutare. Nell'ambito di tale discussione particolare importanza assume la discussione sugli errori medici e sul rischio clinico come misuratori dell'efficacia delle strutture sanitarie: Trucco ne dà uno sguardo d'insieme nel suo lavoro.

Infine, per rispondere a un quesito spesso avanzato sia in chiave teorica che empirica, le metodologie di valutazione dell'efficacia di strutture sanitarie, presentano caratteristiche aspetti simili ma anche importanti differenze rispetto ai più tradizionali studi di tipo biomedico. Il contributo di Battauz, Bellio, Lattuada a riguardo sull'assistenza domiciliare e quello di Rosario, Gregari, Bo, Grassi, Ciccone, Biggeri, Merletti,

Pagano permettono di cogliere tali assonanze e diversità, a partire da casi empirici studiati con strumenti statistici avanzati.

A. METODOLOGIE STATISTICHE DI VALUTAZIONE DELL'EFFICACIA DELLE STRUTTURE SANITARIE

Giorgio Vittadini, Facoltà di Scienze Statistiche Università degli Studi Milano Bicocca

Roberta Carabalona, CRISP

Camillo Rossi, Direzione Generale Sanità Regione Lombardia

1. Efficacia delle strutture sanitarie

Alle analisi di efficienza che considerano il rapporto tra risorse impiegate e risultato di gestione, ai modelli per lo studio della soddisfazione dell'utente, si sono aggiunti dagli anni '90 studi che confrontano l'efficacia di diversi trattamenti o strutture sanitarie 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; Goldstein e Spiegelhalter, 1996).

Il paragone viene effettuato sulla base di opportuni outcome sanitari che misurano l'efficacia. Tuttavia occorre tenere presente che tali outcome sanitari sono influenzati dal "case mix", vale a dire dalla variabilità delle caratteristiche cliniche e socio-demografiche di un certo insieme di pazienti afferenti ad ogni diversa struttura sanitaria (Zaslavsky, 2001). Perciò, sia per effettuare confronti a parità di condizioni (o come spesso si dice *ceteris paribus*), sia per prevenire meccanismi di selezione avversa, in base ai quali le strutture sanitarie potrebbero scegliere di curare preferibilmente o esclusivamente i pazienti meno gravi, occorre "standardizzare" gli outcome sanitari, vale a dire correggerli in modo da tener conto del diverso case mix degli ospedali¹, mediante metodologie denominate di *risk adjustment*, o di *severity adjustment*².

2. Metodologie di Risk Adjustment: standardizzazione diretta e indiretta

Una prima famiglia di procedure di standardizzazione degli outcome può essere denominata come "standardizzazione diretta"³ (Zaslavsky, 2001). L'outcome sanitario "corretto" è dato dalla somma ponderata degli outcome relativi ai diversi strati della popolazione individuata dalle variabili che descrivono il case mix con pesi pari a quelli di ogni strato in una popolazione standard prescelta.

¹ Quali variabili di case mix si citano solitamente età, sesso, numerosità della famiglia, ospedalizzazione precedente, patologia e comorbidità, stato di benessere, stato di salute, stato funzionale, luogo di residenza, stato civile, ampiezza della famiglia (Leary et al., 1998).

² Una metodologia che nel paragonare trattamenti o strutture sanitarie tiene conto delle differenze di caratteristiche dei pazienti (es. morte, funzionalità fisica) si definisce propriamente come "severity adjustment" se il paragone è rispetto ad outcome sanitari, come "risk adjustment" se il paragone è rispetto a costi della cura (Leary et al. 1998). In letteratura usualmente non si distingue tra i due termini e quindi si userà in questa sede il più comune termine di risk adjustment.

³ Si considera ad esempio il caso di studio di parti cesarei in 67 ospedali su 123.850 pazienti "corretti" con standardizzazione diretta. La graduatoria corrispondente, confrontata con quella del modello multilevel, è in questo caso molto simile (Bailit e Garret, 2003). Si consideri ancora ad esempio il caso dell'associazione per le cure mediche dello stato di New York che utilizza la standardizzazione diretta per ogni indicatore Healthcare Cost and Utilization Project (HCUP) relativo a complicanze (Jiang et al., 2001).

Dati: $x_{k\ell}$ = indicatore che definisce il case mix osservato rispetto all' ℓ -esima caratteristica per lo strato i -esimo della popolazione standard $P^{(i)}$ ($i=1, \dots, k$).

y_{kj} = outcome osservato per lo strato k -esimo (rispetto all' ℓ -esima caratteristica) della struttura sanitaria j -esima.

$p_{k\ell}^{(i)}$ = peso relativo dello strato k -esimo individuato dalla caratteristica ℓ -esima della popolazione standard $P^{(i)}$

si ottiene il seguente sistema di pesi :

$$\pi_{k\ell}^{(i)} = \frac{p_{k\ell}^{(i)}}{\sum_i p_{k\ell}^{(i)}} \quad (2.1)$$

L'outcome standardizzato \tilde{Y}_j per la struttura sanitaria j -esima, ($j=1, \dots, q$) è quindi:

$$\tilde{Y}_j = \sum_k \pi_{k\ell}^{(i)} y_{kj} \quad (2.2)$$

La metodologia presenta alcuni problemi: ad esempio non può essere applicata agli outcome che non sono discreti e qualora, a causa di dati mancanti, l'indicatore x_{ℓ} abbia frequenza nulla rispetto ad un determinato strato la scelta della popolazione standard può essere arbitraria; inoltre sul piano operativo risulta arduo ottenere strati con numerosità sufficiente per poter fare confronti (Zaslavsky, 2001);

Una seconda famiglia di metodi di tipo aritmetico è detta della standardizzazione indiretta⁴. (Blangiardo, 1997 e Zaslavsky, 2001). Definiti come in precedenza $x_{\ell i}$, $p_{i\ell}^{(T)}$ e anche

$y_i^{(T)}$ = valore dell'outcome nello strato i -esimo individuato dalla caratteristica x_{ℓ} nella popolazione standard $P^{(T)}$.

p_{ij} = peso relativo dello strato i -esimo individuato dalla caratteristica ℓ -esima per la struttura sanitaria j -esima:

Perciò si ottiene:

$$\hat{Y}_j = \frac{\sum_i y_i^{(T)} * p_{ij}}{\sum_i p_{ij}} \quad (2.3)$$

con \hat{Y}_j outcome standardizzato in modo indiretto per la struttura sanitaria j -esima, ($j=1, \dots, q$)

⁴ Ad esempio nell'ambito del programma di benchmark della Commission for Health Improvement (CHI) del National Health Service (NHS, 2004) si propone la standardizzazione indiretta di alcuni outcome clinici.

Anche la standardizzazione indiretta presenta problemi: permane l'arbitrarietà legata alla scelta della popolazione e la limitazione ad indicatori discreti; inoltre viene ignorato il valore dell'outcome nella struttura sanitaria considerata.

3. Metodologie di risk adjustment: regressione logistica e modelli lineari

I metodi che hanno trovato più consenso sono quelli basati su modelli lineari che permettono di ricavare per ogni ospedale outcome "corretti" al netto degli effetti del diverso case mix.

Nella loro più semplice formulazione, in presenza di outcome dicotomici, si hanno regressioni logistiche del tipo

$$\ln\left(\frac{p_{ij}}{1-p_{ij}}\right) = \alpha + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_k x_{kij} = RS_{ij} \quad (2.5)$$

riesprimibile attraverso opportuni passaggi attraverso la seguente formula

$$p_{ij} = \frac{e^{\alpha + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_k x_{kij}}}{1 + e^{\alpha + \beta_1 x_{1ij} + \beta_2 x_{2ij} + \dots + \beta_k x_{kij}}} = \frac{e^{RS_{ij}}}{1 + e^{RS_{ij}}} \quad (2.6)$$

ove per il paziente i -esimo curato nel j -esimo ospedale, p_{ij} è la probabilità del verificarsi dell'evento favorevole per un outcome sanitario dicotomico, per l'individuo x_1, \dots, x_k sono gli indicatori che descrivono il case mix di caratteristiche del paziente, la quantità a destra dell'uguale in (2.4) RS_{ij} è detta "risk score" (De Long et al., 1997).

Il "risk score atteso" può essere quindi calcolato per ciascun paziente sulla base del valore assunto dalle variabili (X_1, \dots, X_R) in (2.5). Si ottengono, attraverso il modello (2.4) stime per i parametri

$\hat{\beta}_1, \dots, \hat{\beta}_k$ e moltiplicate tali stime per i valori osservati (x_{1ij}, \dots, x_{kij}) nella formula (2.5), si ottiene la probabilità \hat{p}_{ij} del verificarsi dell'evento atteso favorevole per il paziente i -esimo ospitato nella j -esima struttura sanitaria dato il suo case mix e che paragonata con la probabilità dell'evento osservato (0,1), fornisce una stima dell'efficacia dell'ospedale j -esimo sull'individuo i -esimo.

Definito a questo punto il valore atteso degli eventi favorevoli per i pazienti dell'ospedale i -esimo,

$$E(\hat{p}_{ij}) = \sum_j^n \hat{p}_{ij} \quad (2.7)$$

e il corrispondente valore della probabilità degli eventi osservati $\sum(p_{ij})$ il rapporto

$$\frac{E(p_{ij})}{E(\hat{p}_{ij})} \quad (2.8)$$

dà una stima dell'efficacia dell'ospedale j -esimo.

In quest'ottica De Long et al. (1997)⁵ presentano un modello logit a effetti fissi del tipo

$$\ln\left[\frac{p_{ij}}{1-p_{ij}}\right] = \alpha' + \beta' RS_{ij} + \lambda_{ij} w_j \quad (2.9)$$

ove per il paziente i -esimo curato dal j -esimo ospedale w_j è un parametro fisso, RS_{ij} è il risk score, $\exp(\lambda_{ij})$ rappresenta l'"odd ratio" della mortalità⁶ e un modello ad effetti misti:

$$\ln\left[\frac{p_{ij}}{1-p_{ij}}\right] = \alpha' + \beta' RS_{ij} + \delta_j \quad (2.10)$$

⁵ Gli autori considerano un modello logistico di mortalità rispettivamente per l'external benchmark e per l'internal benchmark su 57.187 pazienti sottoposti a Coronary Artery Bypass Graft (CABG) in 30 ospedali dello stato di New York nel periodo dal 1989 al 1992 e su 3654 pazienti di 65 anni e più sottoposti a CABG in 28 differenti ospedali in Iowa ed Alabama, nel periodo 1992-1993.

⁶ Rappresenta la probabilità osservata del verificarsi dell'evento favorevole rispetto alla probabilità attesa tenuto conto del risk score.

ove δ_j è un parametro casuale inerente l'ospedale j -esimo che si distribuisce in modo normale. Gli autori definiscono come: popolazione DEV, la popolazione che ha generato le variabili che descrivono il case mix e il "risk score" RS_{ij} (*development population*); e la popolazione EV, come la popolazione che ha generato gli outcome da standardizzare (*evaluation population*).

Se le due popolazioni non coincidono, per standardizzare i dati della popolazione EV viene usato l'algoritmo sviluppato sui dati della popolazione DEV e tale algoritmo è etichettato dagli autori come *external benchmark*, se invece le due popolazioni coincidono l'algoritmo usato è detto *internal benchmark*⁷. L'importanza del contributo in questione è dato dal fatto che gli autori propongono (De Long et al, 1997) 8 possibili strategie per standardizzare gli outcome sanitari a partire dai modelli (2.8) (2.9)⁸.

Al di là dei lavori sperimentali⁹ il modello logistico, in alcuni casi proposto anche con variabili esplicative inerenti le caratteristiche degli ospedali è il più utilizzato da programmi di benchmark come metodo per ricavare outcome corretti¹⁰, anche in casi in cui la distribuzione degli outcome non sia dicotomica ma politomica ed ordinale. Inoltre, se l'outcome è di tipo quantitativo, si utilizza spesso il modello regressivo, sia in studi teorici (Dubois, 1998; De Long et al., 1997) sia nell'ambito delle grandi banche dati di indicatori¹¹. In questi casi dato un modello regressivo del tipo:

$$y_{ij} = \gamma_q x_{ijq} + e_{ij} \quad (2.10)$$

ove per il paziente i -esimo curato nella j -esima struttura sanitaria y_{ij} è l'outcome sanitario relativo, x_{ijq} è la q -esima variabile esplicativa che descrive il case mix, γ_q il parametro corrispondente, e_{ij} il residuo casuale.

Attraverso (2.10) si stimano i parametri γ_q e i valori "previsti" di y_{ij} dato il case mix.

Viene poi valutata la bontà dell'adattamento e il rapporto tra valori osservati e attesi.

$$E(y_{ij}) / (\hat{y}_{ij}) \quad (2.11)$$

che fornisce una stima dell'efficacia delle cure per j -esimo ospedale.

⁷ Gli autori valutano la capacità dei diversi modelli proposti di classificare correttamente i soggetti sulla base del cosiddetto *c-index*. Tale indice (De Long et al, 1997; Hanley e Mc Neil, 1982), ha un campo di variazione pari a 0.5-1: un valore pari ad 1 indica classificazione perfetta, mentre un valore pari a 0.5 indica classificazione casuale.

⁸ Nel caso 1, utilizzando i coefficienti determinati con l'external benchmark il risk score viene calcolato per ciascun paziente la stima puntuale e intervallare del rapporto tra valori osservati ed attesi al 95. Nel caso 2 vengono corrette le stime calcolate con i dati relativi alla EV con i dati relativi alla DEV. Nel caso 3 ottenuto il RS attraverso il modello (2.9) (2.10) lo si utilizza come predittore in un nuovo modello di regressione. Nel caso 4 e 5 il modello 3 può essere arricchito aggiungendo in un parametro inerente l'ospedale fisso (modello ad effetti fissi caso 4) oppure casuale (modello ad effetti casuali caso 5) I casi 6, 7 ed 8 sono analoghi a quelli indicati come 1, 4 e 5, salvo per il fatto che i risk score sono calcolati sulla base di un algoritmo sviluppato sulla stessa popolazione su cui definiti gli outcome da standardizzare. I modelli ad effetti casuali (5 e 8) portano a stime più conservative, cioè più vicine all'unità, dei modelli ad effetti fissi. I risultati ottenuti analizzando i dati relativi alla mortalità ospedaliera in seguito all'intervento di bypass alle coronarie (3653 pazienti in 28 ospedali), indicano in particolare che il c-index è migliore nel caso in cui il modello comprenda anche l'effetto agente (sia fisso che random) rispetto agli altri casi, a prescindere dal fatto che si sia usato l'external o l'internal benchmark.

⁹ Si consideri ad esempio la proposta di Fortescue et al. (2001) che utilizza la regressione logistica con metodo di "risk adjustment" per patologie di insufficienza renale dopo bypass alle arterie coronarie che richiedono dialisi. Il campione è composto da 9498 pazienti sottoposti a CABG in ospedali accademici di 3° livello.

¹⁰ Il modello logistico è utilizzato come metodo di risk adjustment nei sistemi di indicatori proposti da ORYX- Joint Commission on Accreditation of Healthcare Organisations (JCAHO) (2001), Agency for Healthcare Research Quality (AHRQ) (2003), Canadian Institute for Health Information (CIHI, 2003), Solucient (2003).

¹¹ È il caso dei metodi di benchmark ORYX- JCAHO (2001), AHRQ (2003), CIHI (2003).

4. Metodologie di risk adjustment: modello multilevel

Il modello statistico più utilizzato per studi di valutazione anche in sanità è però il modello multilevel (Hox, 1995) la cui valenza in questo campo è misurata in modo organico al lavoro di Goldstein e Spiegelhalter (1996). Si ha quindi per l' i -esimo paziente curato nell' j -esimo ospedale:

$$y_{ij} = \alpha_q x_{ijq} + u_j + e_{ij} \quad (i=1, \dots, n; \quad j=1, \dots, g, \quad q=1, \dots, p) \quad (2.12)$$

ove y_{ij} è l'outcome relativo all' i -esimo paziente curato nell' j -esima struttura sanitaria; x_{ijq} è la q -esima variabile esplicativa che descrive il case mix dell' i -esimo paziente curato nella j -esima struttura sanitaria; α_q è il parametro ad esso connesso; u_j è una variabile casuale; e_{ij} è il residuo casuale. Si ipotizza che i residui e_{ij} siano impropriamente incorrelati con valore atteso nullo e varianza costante e che ogni variabile casuale u_j con distribuzione $\varphi(u_j)$ con varianza costante descriva l'outcome corretto (o "risk adjusted") al netto dell'influenza del case mix.

Il modello, denominato "a effetti misti"¹², permette di valutare l'efficacia relativa perché, grazie alla struttura gerarchica dei dati, enuclea la variabilità tra e entro gli ospedali, *ceteris 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"¹³. La distribuzione della variabile casuale u_j è ricavata attraverso metodi di inferenza bayesiana: perciò, una volta ipotizzata una distribuzione a priori $\varphi_1(u_j)$ per le variabili casuali u_j si utilizzano anche le informazioni provenienti dal campione per giungere ad una definitiva e più informativa distribuzione a posteriori $\varphi_2(u_j)$ ¹⁴ a partire dalla quale si costruiscono graduatorie di efficacia tra ospedali.¹⁵

Negli altri principali studi finora proposti a riguardo di valutazione dell'efficacia relativa di strutture ospedaliere, mediante il modello multilevel¹⁶, sono state avanzate proposte di generalizzazione sotto il profilo metodologico, in alcuni casi accompagnate da studi empirici.

a) Modello complessivo. In primo luogo, nel caso si ipotizzi che i parametri relativi alle caratteristiche individuali dei pazienti varino casualmente tra ospedale, in luogo del parametro fisso α_q si ha:

$$\alpha_{jq}^* = \alpha_q + f_{jq} \quad (2.13)$$

ove α_q è ancora parametro fisso, f_{jq} è variabile casuale e quindi, α_{jq}^* variabile casuale. In questo caso, in luogo del modello a "effetti misti" si ha il modello a "effetti casuali" (Thomas N., Longford e Rolph 1994):

¹² Poiché i coefficienti α_q relativi alle variabili esplicative x_{ijq} sono fissi mentre i parametri u_j sono casuali.

¹³ Una volta ricavato il valore atteso $E(u_j)$ del parametro casuale u_j lo si corregge mediante il fattore shrinkage nel seguente modo:

$$\hat{u}_j = \frac{n_j s_a^2}{n_j s_a^2 q + s_e^2} E(u_j) \quad [E(u_j) \neq 0]$$

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).

¹⁴ Con Software Bugs Gibbs Sampling basato su un particolare algoritmo Monte Carlo Markov Chain.

¹⁵ 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 $\varphi_2(B)$

¹⁶ Il lavoro di Goldstein e Spiegelhalter (1996) studia l'efficacia di differenti chirurghi in interventi di CABG. Il campione è composto da 87 pazienti in ospedali dello Stato di New York; l'outcome sanitario è il tasso di mortalità intra ospedaliero post intervento distribuito come una Poisson; vi è una sola variabile esplicativa inerente la gravità dei pazienti preoperatoria espressa da punteggi di origine epidemiologica; gli u_j sono relativi a differenti chirurghi; $\varphi_1(u_j)$ e $\varphi_2(u_j)$ sono distribuiti in modo normale

$$y_{ij} = (\alpha_q^* + f_{jq}) x_{ijq} + u_j + e_{ij} \quad (i=1, \dots, n; j=1, \dots, g; q=1, \dots, q) \quad (2.14)$$

Mentre le (2.13) variano tra ospedale ed ospedale per le variabili casuali u_j che fungono da intercette, le (2.15) variano anche per le variabili casuali f_{jq} , che fungono da coefficienti angolari. In quest'ottica Normand et al. (1997) che suggeriscono di utilizzare un modello multilevel logistico con intercetta casuale e struttura gerarchica¹⁷. Il modello è interpretato e risolto nel suo complesso come un modello bayesiano gerarchico che permette di valutare l'efficacia degli erogatori di servizi sanitari basata sulla probabilità a posteriori che gli eventi avversi abbiano una frequenza usuale.

Austin et al. (2002) valutano il grado di concordanza tra graduatorie di ospedali costruite sulla base di quattro diversi metodi di valutazione ottenuti attraverso modelli a effetti casuali e effetti misti e inferenza bayesiana¹⁸ e concludono che tale concordanza non è soddisfacente a parità di outcome.

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 di parametri fissi α_q si ha:¹⁹

$$y_{ij} = (\hat{\alpha}_q + \delta_{jr}) x_{ijq} + u_j + e_{ij} \quad (i=1, \dots, n; j=1, \dots, g; q=1, \dots, p; r=1, \dots, s) \quad (2.15)$$

con $\hat{\alpha}_q$ e δ_{jr} parametro fissi che misurano rispettivamente l'effetto medio dell'ospedale j -esimo e l'effetto interazione fra l'ospedale j -esimo e l' r -esimo sottogruppo di pazienti curati in tale ospedale.

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

$$y_{ijl} = \alpha_q x_{ijq} + \theta_{lj} + e_{ijl} \quad (i=1, \dots, n; j=1, \dots, g; l=1, \dots, m; q=1, \dots, p) \quad (2.16)$$

$$\theta_{lj} = \lambda_j + v_{lj} \quad (2.17)$$

ove θ_{lj} individua l'efficacia associata all' l -esimo reparto del j -esimo ospedale e λ_j l'efficacia dell'ospedale nel suo complesso (Marshall e Spiegelhalter, 2001).

Infine, utilizzando le proprietà teoriche del modello multilivello (Hox, 1995), si può tener conto nell'analisi di efficacia relativa, delle caratteristiche delle diverse strutture ospedaliere z_k ($k=1, \dots, f$). Data l'equazione (2.12), si riesprime la variabile casuale u_j nel seguente modo:

$$u_j = v_k z_{kj} + w_j \quad (j=1, \dots, g; k=1, \dots, f) \quad (2.18)$$

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 ospedaliere²⁰.

¹⁷ Gli autori utilizzano la mortalità a 30 giorni dal ricovero per un campione di 3156 pazienti convenzionati medicare ricoverati per infarto acuto in 96 ospedali.

¹⁸ Gli autori confrontano su un campione di 18850 pazienti ricoverati per infarto acuto al miocardio in 152 ospedali canadesi il *Patient profile method* (Normand et al. 1997); il *Patient profile rank method* (Marshall e Spiegelhalter, 1998); lo *SMR method* (Christiansen e Morris, 1997); lo *SMR rank method* (Marshall e Spiegelhalter, 1998).

¹⁹ Ad esempio Duncan et 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 la variabilità del tasso di mortalità per infarto acuto miocardico tra ospedali scozzesi ognuno dei quali ha efficacia diversa per sottoinsiemi di pazienti raggruppabili in base alle comorbilità.

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 . Si ha in luogo di (10):

$$u_j = v_k z_{kj} + \gamma_{kq} (z_{kj} \bar{x}_{jq}) + w_j \quad (j=1, \dots, g; q=1, \dots, p; k=1, \dots, f) \quad (2.19)$$

con \bar{x}_{jq} media di x_q sull'ospedale j -esimo.

b) Outcome. 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 come nei casi già citati di Normand et al. (1997) descritto in seguito. Se l'indicatore è invece continuo di solito l'indicatore è normale (Morris e Christiansen, 1996; Rice, 2001; Langford e Day, 2001; Marshall e Spiegelhalter, 2001). Numerose sono però le proposte alternative anche in situazioni in cui vi siano outliers (Langford e Lewis, 1988) o dati mancanti tra gli outcome (Leyland e Boddy, 1998). 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), outcome normali di ordine p in uno studio sull'efficacia relativa di ospedali lombardi rispetto a patologie cronic-degenerative (Vittadini e Sanarico, 2002).

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)²¹ 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, 2002).

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):

$$z_{iju} = L_y y_{ij}^+ + \delta_{iju} \quad (i=1, \dots, n; j=1, \dots, g; u=1, \dots, d) \quad (2.20)$$

per poi porre y_{ij}^+ come variabile dipendente nel modello multilevel (2.13) (Goldstein e Leyland, 2001).

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 (Wright e Masters, 1982; Bond e Fox, 2002). Tali proxies sono poi utilizzate come variabili dipendenti in (2.13). 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, 2000).

c) Parametri casuali d'efficacia. Nella quasi totalità delle proposte è assunta distribuzione a priori normale per $\varphi_1(u_j)$. Non mancano però forti critiche per questa scelta (Marshall e Spiegelhalter, 2001), 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

²⁰ Numerosi studi usano in modelli multilevel variabili esplicative inerenti gli ospedali definite a livello aggregato (Dubois et al., 1988; Morris e Christiansen, 1996) e a livello di singolo paziente (Leyland e Boddy, 1998).

²¹ Come nel caso di Mc Culloch et al. (1997) ove si investiga la variabilità tra ospedali della lunghezza della degenza e dell'eventuale nuovo ricovero d'urgenza per pazienti operati.

tengano conto di outliers (Albert e Chib, 1997; Richardson e Green, 1997; Marshall e Spiegelhalter, 2001).²²

d) Confronto fra modelli multilevel e modelli lineari. In tutte le principali proposte di utilizzo del modello multilevel per la valutazione dell'efficacia relativa è costante l'utilizzo di inferenza bayesiana (Thomas N., Longford e Rolph, 1994; Goldstein e Spiegelhalter, 1996; Normand et al., 1997; Morris e Christiansen, 1996; Rice e Leyland, 1996; Leyland e Boddy, 1998; Marshall e Spiegelhalter, 2001). In alcuni studi (Thomas N., Longford e Rolph, 1994; Morris e Christiansen, 1996) si mettono in luce i vantaggi della scelta di una metodologia d'inferenza bayesiana. Tale metodologia 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.

In quest'ottica Morris e Christiansen, 1996 e Normand et al., 1997, confrontano l'utilizzo del modello multilevel e dei modelli regressivi. Nei modelli lineari la precisione delle stime della mortalità attesa può variare molto a causa della variabilità delle dimensioni degli ospedali; non si riesce spesso a individuare con chiarezza tutta l'influenza delle variabili di case mix sull'outcome; può avvenire che il modello regressivo non suddivida in modo adeguato la variabilità campionaria complessiva in variabilità interna alla struttura sanitaria e variabilità sistematica tra le varie strutture; l'uso delle differenze tra valori attesi ed osservati per classificare gli ospedali può portare a distorsioni.

Leyland e Boddy (1998) mostrano invece come la struttura gerarchica dei dati tipica del piano di campionamento del modello multilevel consente di analizzare ciascun paziente nel contesto della struttura sanitaria dove è curato e nell'area geografica dove tale struttura è localizzata²³ e perciò permette di suddividere con molta più chiarezza, rispetto alla regressione logistica, la variabilità degli outcome in variabilità interna e tra le strutture. Infine Austin et al. (2002) mostrano le forti differenze nelle classificazioni di ospedali ottenuto con modello logistico e inferenza frequentista e modello multilevel a effetti casuali e inferenze bayesiane.

5) Scelta degli outcome, delle variabili di case mix

Per quanto un metodo di risk adjustment compensi l'influenza sugli outcome del differente case mix di ogni paziente, non sarà possibile annullare tale influenza (CIHI, 2003). Molto dipenderà anche dall'opportuna scelta di outcome e variabili di case mix, oltre che del metodo di un risk adjustment prescelto. Tuttavia la scelta di outcome e variabili case mix, oltre che dagli scopi dell'analisi è condizionata dai dati a disposizione e dalla tipologia della loro fonte.²⁴

Dalla letteratura e dall'evidenza empirica vengono alcuni risultati:

a) Mortalità intraospedaliera. Il più diffuso indicatore di efficacia è la mortalità o il tasso di mortalità intra-ospedaliera, classificabile come outcome clinico. Tale outcome è il più usato sia in studi sperimentali sulla valutazione condotti su piccoli campioni (Thomas N., Longford e Rolph,

²² Oppure si è proposto di partire da distribuzioni a priori non informative $\varphi_1(u_j)$ che, grazie all'informazione ricavata dalla distribuzione del campione, permettano di giungere a distribuzioni a posteriori informative per $\varphi_2(u_j)$ (Vittadini e Sanarico, 2002).

²³ Gli autori studiano la mortalità per infarto: in casi di decesso non ospedalieri, al fine di evidenziare l'effetto "area geografica e decessi in ospedale oppure entro 30 giorni dalla dimissione, allo scopo di analizzare anche l'effetto "ospedale".

²⁴ La creazione (Mc Glynn et al., 1998) di un sistema informativo adeguato è alla base del processo di monitoraggio della qualità dell'assistenza sanitaria.

1994; Leyland e Boddy, 1998; Normand et al., 1997; Goldstein e Spiegelhalter, 1996; Marshall e Spiegelhalter, 1999; Marshall e Spiegelhalter, 2001) sia in tutti i grandi progetti di benchmark perchè sembra intrinsecamente legato all'efficacia e rispetto agli altri outcome clinici, è più facilmente rilevabile.

In relazione al tasso di mortalità sono stati proposti numerosi indici di severità che vogliono quantificare a priori il rischio di morte per un paziente affetto da determinate patologie. Molte delle numerosissime misure di severità sono generate in ambiti commerciali o di politica sanitaria (Iezzoni et al., 1996); non mancano però proposte che quantificano il rischio di morte assegnando modalità ordinali o probabilità alle diverse condizioni patologiche a partire da rigorosi studi scientifici. Numerosi sono stati anche i tentativi di classificazione di tali indici (Thomas N., Longford e Rolph, 1994; Iezzoni, 1997²⁵; Leary et al., 1998) alcuni di origine clinica²⁶, altri su schede di dimissione rielaborate su dati di origine clinica²⁷. In mancanza di tali indici si può correggere il tasso di mortalità con variabili di case mix²⁸.

Non pochi autori ritengono comunque che la mortalità, anche se ricavata da dati di tipo clinico sia insufficiente per descrivere l'esito della cura dato il suo carattere esclusivamente dicotomico che prevede solo la possibilità di sopravvivenza o meno di un paziente dopo la cura (Goldstein e Spiegelhalter, 1996; Marshall e Spiegelhalter, 2001)²⁹. Anche la "correzione" della mortalità con indici di severità non genera misure corrette capaci da sole di spiegare la differente qualità degli ospedali³⁰ (Iezzoni et al., 1996).

Anche fonti amministrative forniscono con sistematicità nel tempo e nello spazio informazioni a riguardo del tasso di mortalità intraospedaliera a 30 giorni non solo per stati e regioni, ma anche per singole strutture sanitarie³¹ con relativa correzione per case mix. In caso di dati sulla mortalità ricavati da fonte amministrativa, diventa ancor più stringente la critica al tasso di mortalità sia pur corretto, come unico misuratore dell'efficacia delle strutture sanitarie (Thomas N., Longford e Rolph, 1994)³².

b) Outcome di tipo epidemiologico e clinico³³. I dati epidemiologici sono originati nell'ambito di lavori di ricerca allo scopo di verificare cause all'origine di patologie, effetti di farmaci o trattamenti medici all'interno di disegni sperimentali o quasi sperimentali ove, accanto a campioni di persone

²⁵ Gli autori analizzano in due successivi studi diverse proposte di indici di severità mediante funzione logistica. La prima analisi empirica è effettuata su 11.880 pazienti di 100 ospedali, ricoverati per infarto acuto al miocardio e la mortalità considerata è quella intra-ospedaliera. Nella seconda analisi Iezzoni, (1997) considera anche 7765 pazienti con CABG, 18016 con polmonite, 9407 con ictus.

²⁶ Alcuni indici sono espressi in termini di probabilità come il *CHQC* (Sirio et al., 1994) basato su specifiche variabili cliniche; i *Medis Groups* in versione empirica (Steen et al., 1993); il *Cardiac Surgery Reporting System* (Chassin et al., 1996); il *Northern New England Cardiovascular Disease Study Group* (O' Connor et al., 1992); il *Medical Mortality Predictor Sistem*, (Steen et al., 1993). Altri indici sono espressi attraverso punteggi come i *Medis Groups* (Blumberg, 1991); i *Physiology Score 1 e 2* (Knaus et al., 1985; Knaus et al., 1991).

²⁷ Il *Disease staging (DS)* e il *DS-Mortality probabilità Stage* (Markson et al., 1991); il *Patient Management Categories (PMCs) Severity Score* (Young, Kholer, Kowalski, 1994); il "*Comorbidity index*" (Charlson et al., 1987).

²⁸ È ciò che avviene nei programmi di benchmark già citati.

²⁹ Addirittura è considerato a volte causa di dimissioni ingiustificate e premature per evitare che la mortalità verificatasi prima del 30° giorno influenzino negativamente la valutazione dell'ospedale (Goldstein e Spiegelhalter, 1996).

³⁰ Iezzoni (1997) mostra che differenti misure di severità correggono in modo differente il tasso di mortalità e che comunque il tasso di mortalità corretto non spiega la differente efficacia degli ospedali.

³¹ Esistono raccolte di tassi di mortalità "corretti" mediante modelli regressivi o logistici per ospedali quali quelle dell' "Health Care Finance Administrative" negli Stati Uniti (Daley et al., 1988), 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).

³² Thomas N., Longford e Rolph (1994) affermano in particolare che le correzioni effettuabili sui tassi di mortalità per tener conto del case mix dei pazienti rimangono comunque troppo limitate.

³³ Ad esempio Morris e Christiansen, (1996) utilizzano le frequenze relative dei rigetti in trapianti di rene; in Italia, Rizzi et al. (1999) studiano il tasso di incidenza endemico delle infezioni postoperatorie e nelle lesioni da decubito la variabile ordinale di Norton.

sottoposte a trattamenti o farmaci opportunamente selezionati, vi sono gruppi di controllo composti da persone dalle caratteristiche simili ma non sottoposti a trattamenti. I dati di origine clinica (Damberg et al., 1998) sono dati che descrivono gli effetti delle patologie e della cura sul paziente e sono generati dal medico durante la cura del paziente (mediante cartella clinica o altro documento). In alcuni casi sono proposte misure di severità con outcome diverso dalla mortalità (Leary et al., 1998)³⁴. Evidentemente gli outcome e le variabili di case mix di origine epidemiologica e clinica sono quelle che insieme agli outcome clinici individuano più direttamente l'efficacia delle istituzioni sanitarie (Kerr ed. al., 1998) in quanto consentono di riportare la diagnosi con maggiore accuratezza³⁵.

Tuttavia per ciò che concerne i dati epidemiologici occorre dire che è estremamente difficile estendere il significato di studi sulla valutazione condotti su piccoli campioni a casi più generali per valutare istituzioni sanitarie con regolarità nel tempo. Motivi di costo, tempo, rigore scientifico, necessità di reperire personale atto allo scopo, sistematicità rendono arduo utilizzare sistematicamente tali dati. Per ciò che concerne i dati di origine clinica, d'altra parte, rispetto agli outcome, si possono incontrare numerosi problemi di identificazione dei pazienti di standardizzazione, di oggettività di valutazione medica, rispetto agli indici di severità che possono essere insufficienti a cogliere a pieno il case mix dei pazienti³⁶.

c) Proxy di indicatori di efficacia. Una proposta alternativa è quella di definire proxy degli outcome clinici, che descrivono l'uso di particolari tecniche e procedure il cui utilizzo è condizione necessaria anche se non sufficiente per l'ottenimento di risultati più efficaci. Tali proxy possono essere individuate molto più facilmente di quanto avvenga per gli outcome clinici. A questo tipo di outcome appartengono ad esempio indicatori di volume di processi connessi con l'efficacia nei progetti di *benchmark*. La fonte di tali proxy è solitamente dato da indagini ad hoc sulle strutture sanitarie³⁷. La non sistematicità di tali indagini, il ritardo nella pubblicazione dei risultati, di tempo e di costo ostacolano la loro introduzione generalizzata (Damberg et al., 1998).

d) Eventi Sentinella. Una quarta alternativa è quella di rilevare "eventi sentinella"³⁸. Gli eventi sentinella riconducibili a episodi particolari più facilmente individuabili rispetto agli outcome clinici possono essere ricavati: da registri particolari riguardanti malformazioni congenite, indagini su cause di morte particolare³⁹; in casi particolari anche da fonti amministrative⁴⁰; in altri

³⁴ Il *Computer Severity Index* (Iezzoni e Daley 1992); l'*Asthma severity index*", (Steinwachs et al., 1995); il *Type II Diabetes Patient Outcomes Research Team* (Greenfield et al., 1995).

³⁵ Vedi a proposito il manuale della JCAHO, (1998).

³⁶ Goldstein e Spiegelhalter, (1996) affermano che indicatori di gravità delle malattie possono non tener conto dell'influsso di variabili sociali e altri fattori; Epstein e Schneider, (1997) afferma che la pubblicazione del rapporto relativo al tasso di rischio per pazienti che dovevano sottoporsi a by pass coronarico ha portato a una sopravvalutazione di tale rischio. Leyland e Boddy, (1998) e Marshall e Spiegelhalter, (1999) osservano come gli indicatori di gravità proposti non siano sufficienti a descrivere il differente stato di salute dei pazienti prima della cura. Anche l'indice di gravità proposto dal "The Prospective Payment System Rand Quality and Care Study" è criticato in quanto spiega molto poco la variabilità della mortalità (Jenks et al., 1988).

³⁷ I programmi Oryx-Jcaho (2001), CIHI (2003), National Health Service (NHS)(2004) svolgono indagini ad hoc per raccogliere indicatori clinici.

³⁸ Sono eventi insospettati 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, 1998).

³⁹ Ad esempio in Friuli sono disponibili banche dati che forniscono informazioni a livello del singolo paziente a riguardo di anagrafe assistiti, vaccinazioni, prescrizioni farmaceutiche, referti di anatomia patologica, prestazioni ambulatoriali.

⁴⁰ E' il caso ad esempio delle complicanze ricavabili dalle fonti amministrative del programma Medicare (Solucient, 2003).

casi ancora da rilevazioni ad hoc⁴¹. Tuttavia sono un indicatore più indiretto degli outcome clinici e la qualità del dato dipende intrinsecamente dalla qualità della fonte da cui è ricavato.

e) Outcome sulla qualità della vita. Un'altra alternativa agli outcome clinici è data dall'utilizzo di outcome inerenti la qualità della vita, che descrivono l'esito di cure in termini di benessere generale e particolare, funzionalità fisica e psicologica. Tali outcome sono rilevabili in modo più semplice degli outcome clinici su scala ordinale attraverso opportuni questionari; tuttavia ognuno di loro è specifico per gruppi di particolari patologie e la loro rilevazione può avvenire solo attraverso indagini ad hoc con i problemi già ricordati.

Si pensi ad esempio all'SF-12 (Ware, 1993) particolarmente atto a rilevare la qualità della vita in caso di malattie croniche, o dell' FIM (Tesio et al., 1997) valido per la disabilità.

f) Outcome di struttura. Infine vi sono numerosi outcome quali i trasferimenti tra strutture, i ricoveri ripetuti, i ritorni in sala operatoria che non riguardano né le condizioni cliniche dei pazienti, né la loro funzionalità, fisica e mentale, né eventi sentinella, ma solo condizioni necessarie perché si verificino conseguenze negative per i pazienti. Tali outcome, che si possono denominare in generale come "outcome di struttura"⁴², sono senz'altro più imprecisi in quanto mettono in luce solo condizioni necessarie per l'inefficacia ma sono anche più facilmente rilevabili in modo sistematico da dati amministrativi (Damberg et al., 1998). Tali dati sono generati dai datori di lavoro e dagli individui stessi, nel momento in cui un paziente si rivolge ad un medico/laboratorio o ad un ospedale e vengono utilizzati dallo stato e dalle compagnie assicurative e da altri enti preposti al pagamento delle spese sanitarie. In connessione con outcome ricavabili da dati amministrativi si possono utilizzare indicatori di case mix di tipo generale (età, sesso, tipo di ricovero e dimissioni) o specifiche per patologie inerenti ad esempio comorbidità⁴³ e indici di severità anch'essi ricavabili da dati amministrativi⁴⁴. Inoltre si utilizzano quali variabili di case mix anche caratteristiche della struttura sanitaria.⁴⁵ I dati amministrativi, essendo generalmente su base informatizzata, permettono di raccogliere con sistematicità grandi quantità di dati e sono più facili da maneggiare e più sistematici nel tempo e nello spazio. Tuttavia, oltre a dare una informazione solo indiretta sull'efficacia, determinano altri problemi (Damberg et al., 1998) poiché possono: mancare di informazioni relativamente a fondamentali caratteristiche cliniche, presentare variabilità nella codifica delle diagnosi e delle procedure, mancare di accuratezza e completezza nella codifica,

⁴¹ Come in Oryx-Jcaho (2001), CIHI (2003), AHRQ(2003).

⁴² È ancora il caso del Friuli che rileva nella sua banca dati appropriatezza dei ricoveri, informazioni chirurgiche e post operatorie, tempo di ricovero.

⁴³ Tali variabili sono utilizzate oltre che dai grandi programmi di benchmark menzionati in numerosi studi scientifici. Rice e Leyland, (1996) in uno studio inerente sui 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 effettuato sulla base del tasso di mortalità per infarto miocardico propongono di usare quali variabili individuali provenienti dal General Register Office for Scotland: età, sesso, stato civile tipo di ricovero, diagnosi secondaria, eventuale intervento chirurgico precedente. Marshall e Spiegelhalter (1999) sottolineano la necessità di usare un vasto insieme di indicatori quali diagnosi, sesso, età e gravità della malattia per superare le differenze dovute alla diversa condizione del paziente prima della cura.

⁴⁴ Si considerino ad esempio gli *All Patients Refined Diagnosis Related Groups (APR – DRG)*, i *Refined Diagnosis Related Groups (R-DRG)* l'*Acuity Index Method (APR DRG)*, 1993, 1994).

⁴⁵ Silber et al. (1995), in uno studio condotto nel 1990-1991 su 73174 pazienti distribuiti in fra 137 ospedali mediante modello logit ,mostra che la variazione di outcome clinici dall'esito avverso fra ospedali è maggiormente spiegata dai variabili esplicative inerenti pazienti piuttosto che da variabili relative ad ospedali quali numero letti, dotazione di attrezzature e volume e tipologia di attività Vittadini e Sanarico (2002) in uno studio su 65000 pazienti chirurgici di ospedali pubblici lombardi mediante il modello Multilevel con variabili esplicative relative a pazienti e ospedali mostrano l'importanza che le dimensioni degli ospedali contribuiscono in modo significativo a spiegare la variabilità di outcome di struttura.

essere caratterizzati da *upcoding* (vale a dire una codifica effettuata per ottenere un rimborso maggiore), l'incertezza sulla scansione temporale degli eventi codificati⁴⁶.

Bibliografia

1. 3M HEALTH INFORMATION SYSTEMS, (1993), *All Patient Refined Diagnosis Related Groups. Definition Manual*, Wallingford, Conn.
2. AALEN O., (1989), *A linear regression model for the analysis of life times*, *Statistics in Medicine*, 8: 907-925.
3. AALEN O., (1993), *Further results on the non-parametric linear regression model in survival analysis*, *Statistics in Medicine*, 12: 1569-1588.
4. AALEN O., BORGAN O., AND FAKJEAR H., (2001), *Covariate adjustment of event histories estimated from Markov chain: the additive approach* *Biometrics*, 57: 993-1001.
5. ABRARIA C., EMANUELE N., COLWELL J., HENDERSON W., COMSTOCK J., LEVIN S., NUTTALL F., SAWIN C., (1992), *For VA Cooperative Study Group. Glycemic control and complications in type II diabetes: design of a feasibility trial*, *Diabetes Care*, 15: 1560-71.
6. AHRQ AGENCY FOR HEALTHCARE RESEARCH AND QUALITY, (2003), *Guide to Inpatient Quality Indicators* Rockville, MD, US Department of Health and Human Services, <http://www.ahrq.gov/data/hcup>
7. ALBERT J, CHIB S. (1997) Bayesian tests and model diagnostics in conditionally independent hierarchical models, *Journal of the American Statistical Association*, 92, 916-25
8. ANDERSON C., RUBENACH S., MHURCHU C. N., CLARK M., SPENCER C., WINSOR A., (2000). *Home or hospital for stroke rehabilitation? Results of a randomised controlled trial. I: health outcomes at 6 months*, *Stroke*, 31: 1024-1031.
9. ANDREWS L.B. et al., (1997), *An alternative strategy for studying adverse events in medical care*, *The Lancet*, 349.9048, pp. 309-13.
10. AUSTIN P.C. (2002), *A comparison of bayesian methods for profiling hospital performance*. *Med Decis Making*; 22: 163-72.
11. BAILIT J., GARRET J., (2003), *Comparison of risk adjustment methodologies for cesarean delivery rates*, *Obstet. Gynecol. Jul.*, 102 (1) , p. 45-51.
12. BASINI V. et al., 2002, *FMEA-FMECA Analisi dei modi di errore/guasto e dei loro effetti nelle organizzazioni sanitarie*, Dossier 75, Regione Emilia-Romagna/Agenzia sanitaria regionale, www.regione.emilia-romagna.it.

⁴⁶ Numerosi altri problemi tipici di ogni indagine statistica (tipo di campionamento, numerosità complessiva, problema di *data mining*, errori dovuti al processo di selezione dei soggetti, risposte dei soggetti sovra-sotto stimate, non risposta, dimissioni non casuali, dropout in studi longitudinali, ecc.) vengono segnalati in letteratura (Mc Glynn e Escance, 1998).

13. BATES D.W. et al., (1995), *Relationship between medication errors and adverse drug events*, J Gen Intern Med, Vol. 10, pp. 199.
14. BATES D.W. et al., (1998), *Effect of Computerized Physician Order Entry and a team intervention on prevention of serious medication errors*, JAMA, Vol. 280, n. 15, pp. 1311-1316.
15. BECKMANN U. et al., (1996), *The Australian incident monitoring study in intensive care: AIMS-ICU. An analysis of the first year of reporting*, Anaesth Intens Care, Vol. 24, pp. 314-19.
16. BLANGIARDO G.C. (1997), *Elementi di demografia*, Il Mulino, Bologna.
17. BLUE L., LANG L., MCMURRAY J. J. V., DAVIE A. P., MCDONAGH T. A., MURDOCH D. R., PETRIE M. C., CONNOLLY E., NORRIE J., ROUND C. E., FORD I., MORRISON C. E., (2001), *Randomised controlled trial of specialist nurse intervention in heart failure*, British Medical Journal, 323: 715-718.
18. BLUMBERG M.S., (1991), *Biased estimates of expected acute myocardial infarction mortality using Medis Groups admission severity groups*, JAMA, 265, 2965-2970.
19. BOND T.G., (2001), *Fox CM Applying the Rasch Model. Fundamental Measurement in the Human Sciences.*, Erlbaum Publisher, London.
20. BOND T.G., FOX C.M., (2002), *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, Lawrence Erlbaum Associates, Publishers, Mahwah, New Jersey.
21. BONNEMA J., ANNEKE VAN WERSCH M E.A., VAN GEEL A.N., PRUYN J.F.A., SCHMITZ P.I.M., PAUL M.A., WIGGERS T., (1998), *Medical and psychosocial effects of early discharge after surgery for breast cancer: randomised trial*, British Medical Journal, 316: 1267-71.
22. BRENNAN T.A. et al., (1991), *Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I*, N Engl J Med, Vol. 324, pp. 370-76.
23. BRENNAN T.A. et al., (1991), *The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II*, N Engl J Med, Vol. 324, pp. 377-84.
24. CANADIAN INSTITUTE FOR HEALTH INFORMATION, (2003), *Hospital report 2002, Acute care Technical Summary*, <http://secure.cihi.ca/cihiweb/splash.html>.
25. CENSIS, (2001), *Dossier: rischi ed errori nella sanità italiana. La gestione delle responsabilità nelle politiche assicurative per la salute*, pp. 1-54, www.censis.it.
26. CHARLSON M.E., POMPEI P, ALES K.L, MACKENZIE C.R., (1987), *A new method of classifying prognostic comorbidity in longitudinal studies: development and validation*, J Chronic Dis., 40: 373-383.
27. CHASSIN M.R., HANNAN E.L., DE BUONO B.A., (1996), *Benefits and hazards of reporting medical outcomes publicly*. N Engl J Med. 8;334(6):394-8.

28. CHRISTIANSEN C.L., MORRIS C.N., (1997), *Improving the statistical approach to health care provider profiling*, *Ann Intern Med.*; 127: 764-8.
29. CLARKE P., GRAY A., LEGOOD R., BRIGGS A., HOLMAN R., (2003), *The impact of diabetes-related complications on healthcare costs: results from the United Kingdom Prospective Diabetes Study*, (UKPDS Study N° 65) *Diabetic Medicine* 20: 442-450.
30. D'AGOSTINO R. B., (1998), *Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group*, *Statistics in Medicine*, 17: 2265-2281.
31. DALEY J., JENCKS S., DRAPER D., LENHART G., THOMAS N, WALKER J., (1988), *Predicting Hospital-Associated Mortality for Medicare Patients. A Method for Patients with Stroke, Pneumonia, Acute Myocardial Infarction, ND Congestive Heart Failure*, *JAMA*, 269 (24): 3617-24.
32. DAMBERG C., KERR E.A., MC GLYNN E.A., (1998), *Description of Data Sources and Restricted issue*, in MC GLYNN E.A., ESCANCE J. (EDS.), *Health Information Systems. Design Issues and Analytic Application*, RAND Health Corporation , cap. 5, 43-76.
33. DAVIES L., WILKINSON M., BONNER S., CALVERLEY P.M.A., ANGUS R.M., (2000), *Hospital at home versus hospital care in patients with exacerbations of chronic obstructive pulmonary disease: prospective randomised controlled trial*, *British Medical Journal*, 321: 1265-1268.
34. DE ROSIER J. et al., (2002), *Using Health Care Failure Mode and Effect Analysis: the VA National Center for Patient Safety's Prospective Risk Analysis System*, *Journal on Quality Improvement*, Vol. 28, n. 5, pp. 248-65.
35. DE LONG E.R., PETERSON E.D., DE LONG D.M., MUHLBAIER L.H., HACKETT S. AND MARK D.B. (1997), *Comparing risk-adjustment methods for provider profiling*, *Stat Med*, 16: 2645-64.
36. DONABEDIAN A., (1990), *La qualità dell'assistenza sanitaria. Principi e metodologie di valutazione*, La Nuova Italia Scientifica, Roma.
37. DONNAN P.T., LEESE G.P., MORRIS A.D., (2000), *Hospitalizations for people with type 1 and type 2 diabetes compared with the non diabetic population of Tayside, Scotland: a retrospective cohort study of resource use*. DARTS/MEMO Collaboration. The Diabetes Audit and Research in Tayside, Scotland/ The Medicine Monitoring Unit Collaboration *Diabetes Care*, 23: 1774-9.
38. DRAFT FEDERAL INFORMATION PROCESSING STANDARDS PUBLICATION 183, (1993), *Integration Definition for Function modeling (Idef)*, www.idef.com.
39. DUBOIS R.W., BROOK R.H., ROGERS W.H., (1988), *Adjusted hospital death rates: A potential screen for quality of medical care*, *American Journal of Public Health* 77, 1162-7.
40. DUNCAN C., JONES K., MOON G., (1998), *Context, composition and heterogeneity: using multilevel models in health research*, *Social Science and Medicine* 46, 97-117.

41. ELKAN R., KENDRICK D., DEWEY M., HEWITT M., ROBINSON J., BLAIR M., WILLIAMS D., BRUMMELL K., (2001), *Effectiveness of home based support for older people: systematic review and meta-analysis*, British Medical Journal, 323: 1-9.
42. EPSTEIN A., SCHNEIDER F.C., (1997), *Influence of cardiac-surgery performance reports on referral practices and access to care*, New England Journal of Medicine, 335, 251-6.
43. FETTERS M.D. et al., (1997), *Adverse events in primary care identified from a Risk-management Database*, J Farm Pract, Vol. 45, pp. 40-46.
44. FOLLI H.L. et al., (1987), *Medication error prevention by clinical pharmacists in two children's hospitals*, Pediatrics, Vol. 79, pp. 718-22.
45. [FORTESCUE E.B., KAHN K., BATES D.W.](#), (2001), *Development and validation of a clinical prediction rule for major adverse outcomes in coronary bypass grafting*. Am J Cardiol. 1; 88 (11): 1251-8.
46. GAGLIARDINO J.J., OLIVERA E.M., BARRAGAN H., PUPPO R.A., (1993), *A simple economic evaluation model for selecting diabetes health care strategies*. Diabetic Medicine 10: 351-4.
47. GOLDSTEIN H, LEYLAND A.H., (2001), *Further developments* in GOLDSTEIN H., LEYLAND A.H. (eds.) *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 175-186.
48. GOLDSTEIN H, SPIEGELHALTER D.J., (1996), *League tables and their limitations: statistical issues in comparison of institutional performance*, JRSS, Ser. A, 159, 3, pp. 385-443.
49. GOLDSTEIN H., (2000), *Random effects meta-analysis of trials with binary outcomes using multilevel models*, Multilevel Modelling Newsletter, vol. 11, n. 1.
50. GRAY A., RAIKOU U., MCGUIRE A. et al., (2000), *Cost effectiveness of an intensive blood glucose control policy in patients with type 2 diabetes: economic analysis alongside randomised controlled trial*, British Medical Journal, 320: 1373-1378.
51. GREENFIELD, A., L. SULLIVAN, K. A. DUKES, ET AL., (1995) *Development and Testing of a New Measure of Case Mix for Use in Office Practice*, Med Care, 33 (4): AS 47-AS 55.
52. HANLEY J.A. AND MCNEIL B.J., (1982), *The meaning and use of the Area under a Receiver Operating Characteristic (ROC) Curve*, Radiology, 143: 29-36.
53. HEALY M.J.R., (2001), *Multilevel Data and Their Analysis* in GOLDSTEIN H., LEYLAND A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 1-12.
54. HENRIKSSON F., AGARDH CD., BERNE C., BOLINDER J., LONNQVIST F., STENSTROM P., OSTENSON G., JONSSON B., (2000), *Direct medical costs for patients with type 2 diabetes*, Sweden Journal of Internal Medicine, 248: 387-96,
55. HOX J.J., (1995), *Applied Multilevel Analysis*, TT-Publikaties, Amsterdam,
56. <http://www.jcaho.com/accredited+organizations/behavioral+health+care/oryx/index.htm>.
57. IEZZONI L.I. E DALEY J., (1992), *A description and Clinical assessment of the computerized severity index*, Quality review bulletin, 44-52.

58. IEZZONI L.I., ASH S.A., SCHWARTZ M., DALEY J., HUGHES J.S. AND MACKIERNAN Y.D., (1996), *Judging hospitals by severity-adjusted mortality rates: the influence of the severity-adjustment method*, Am J Public Health.; 86: 1379-87.
59. IEZZONI L.I., (1997), *The Risks of Risk Adjustment*, JAMA; 278 (19): 1600-7.
60. JACHO, (2002), *Joint Commission Perspectives on Patient Safety*, Joint Commission Resources Publication, Vol. 2, Issue 4.
61. JANG H.J., STRYER D., (2003), *Multiple hospitalizations for patients with diabetes*, Diabetes Care, 26: 1421-26.
62. JCAHO JOINT COMMISSION ON ACCREDITATION OF HEALTHCARE ORGANIZATIONS, (2001), *Oryx Technical Implementation Manual*.
63. JENCKS S.F., DALEY J., DRAPER D., THOMS N., LENHART G., WALKER J., (1988), *Interpreting Hospital Mortality Data. The Role of Clinical Risk Adjustment*, Journal of the American Medical Association 260, 3611-16.
64. JIANG H.J., CICCONE K., URLAUB C.J., BOYD D., MEEKS G., HORTON L., (2001) *Adapting the HCUP QIs for hospital use: the experience in New York State*, Healthcare Association of New York State, Albany, New York State.
65. JOINT COMMISSION ON ACCREDITATION OF HEALTHCARE ORGANIZATIONS, (1998), *National Library of Healthcare Indicators*, Health Plan and Network Edition.
66. KASS B.L., (2001), *Reducing and preventing adverse drug events to decrease hospital costs*, Research in Action, Issue 1, AHRQ Publication Number 01-0020, Rockville, MD, March.
67. KERR E.A. MC GLYNN E.A., DAMBERG C., (1999), *Developing and using a clinical information system*, in MC GLYNN E.A., DAMBERG C.L., KERR E.A., BROOK R.H. (eds.), (1998), *Health Information Systems. Design Issues and Analytic Application*, RAND Health Corporation, cap. 7, 141-187.
68. KING H., AUBERT RE., HERMANN WH., (1998), *Global burden of diabetes, 1995-2025. Prevalence numerical estimates and projections*, Diabetes Care, 21: 1414, 31.
69. KLEIN M.L., (1997), *Moeschberger. Survival analysis*, Springer, Verlag.
70. KNAUS W.A, DRAPER E.A., WAGNER D.P., ZIMMERMANN J.E., (1985), *APACHE II: a severity of disease classification system*, Crit. Care Med., 13: 818-829.
71. KNAUS W.A., WAGNER D.P., DRAPER E.A., et al., (1991), *The APACHE III prognostic system: risk prediction of hospital mortality for critically ill hospitalised adults*, Chest.,100: 1619-1636.
72. KOHN L.T. et al., INSTITUTE OF MEDICINE (U.S.) EDITORS, (1999), *To err is human: building a safer health system*, Committee on Quality of Health Care in America, National Academy Press, I-XXIV, pp.1-85, www.nap.edu.

73. LANGFORD I., DAY R.J., (2001), *Poisson Regression* in GOLDSTEIN H., LEYLAND A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, 45-58.
74. LANGFORD I., LEWIS T., (1988), *Outliers in multilevel data*, *Journal of the Royal Statistical Society A* 161, 121-60.
75. LEAPE L.L. et al., (1993), *Preventing medical injury*, *Qual Rev Bull*, Vol. 8, pp. 144-49.
76. LEAPE L.L. et al., (1995), *Systems Analysis of adverse drug events*, *JAMA*, Vol. 274, n. 1, pp.35-43.
77. LEAPE L.L., (1994), *Error in medicine*, *JAMA*, Vol. 272, n. 23, pp. 1851-57, December.
78. LEAPE L.L., (2002), *Reporting of Adverse Events*, *N Engl J Med*, Vol. 347, n. 20, pp. 1633-38.
79. LEARY J.O., KEELER E.B., DAMBERG C., KERR E.A., *An overview of risk adjustment*, in Eds. Mc GLYNN E.A., DAMBERG C.L., KERR E.A., BROOK R.H., (1998) *Health Information Systems. Design Issues and Analytic Application*. RAND Health Corporation, cap. 9, 213-249.
80. LESAR T.S. et al., (1997), *Factors Related to Errors in Medication Prescribing*, *JAMA*, Vol. 277, n. 4, pp. 312-17.
81. LEYLAND A.H., BODDY F.A., (1998), *League tables and acute myocardial infarction*, *Lancet* 351, 555-8.
82. MARKSON L.E., NASH D.B., LOUIS D.Z., GONNELLA J.S., (1991), *Clinical outcomes management and disease staging*. *Eval Health Professions*, 14: 201-227.
83. MARSHALL E.C., SPIEGELHALTER D.J., (1998), *Reliability of league tables of in vitro fertilization clinics: retrospective analysis of live birth rates*, *Br Med J* 1998, 316: 1701-1705.
84. MARSHALL E.C., SPIEGELHALTER D.J., (1999), *League tables of in vitro fertilisation clinics: How confident can we be about the rankings?*, *British Medical Journal*, 317, 1701-04.
85. MARSHALL E.C., SPIEGELHALTER D.J., (2001), *Institutional Performance*, in GOLDSTEIN H., LEYLAND A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 127-142.
86. MC CULLOCH C., LEYLAND A.H. BODDY F.A., (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.
87. MC GLYNN E.A., DAMBERG C.L., KERR E.A., BROOK R.H., (1998), *Health Information Systems. Design Issues and Analytic Application*, RAND Health Corporation.
88. MC GLYNN E.A., ESCANCE J., (1998) *Health Information Systems. Design Issues and Analytic Application*, RAND Health Corporation.
89. MILLER S.C., GOZALO P., MOR V., (2001), *Hospice enrollment and hospitalization of dying nursing home patients*, *The American Journal of Medicine*, 111: 38-44.

90. MORRIS C.N., CHRISTIANSEN C.L., (1996), *Hierarchical models for ranking and for identifying extremes, with applications*, in BERNARDO J.O., BERGER J.O., DAWID, A.P., AFM (eds.) *Bayesian Statistics 5*, Oxford University Press, 277-97.
91. MOSS S.E., KLEIN R. AND KLEIN B., (1999) *Risks factors for hospitalization in people with diabetes*, *Archives of Internal Medicine*, 159: 2053-7.
92. MUNARI L., (2002), *L'assistenza su misura. Orientarsi fra gli strumenti valutazione delle attività sanitarie*, Centro Scientifico Editore, Torino.
93. NAYLOR M.D., BROOTEN D., CAMPBELL R., JACOBSEN B.S., MEZEY M.D., PAULY M.V., SCHWARTZ J.S., (1992), *Comprehensive discharge planning and home follow-up of hospitalized elders*, *Journal of the American Medical Association*, 281: 613-620.
94. NHS NATIONAL HEALTH SERVICE CHI COMMISSION FOR HEALTH IMPROVEMENT, (2004), *A commentary on star ratings 2002- 2003*, <http://www.chi.nhs.uk/ratings>
95. NORMAND S.L., GLICKMAN M.E. AND GATSONIS C.A., (1997), *Statistical Methods for Profiling Providers of Medical Care: Issues and Application*, *Journal of the American Statistical Association*, 92 (439): 803-14.
96. O'CONNOR G.T., PLUME S.K., OLMSTEAD E.M., et al., (1992), *Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery*. Northern New England Cardiovascular Disease Study Group. *Circulation*, 85 (6): 2110-8.
97. PHILLIPS J. et al., (2001), *Retrospective analysis of mortalities associated with medication errors*, *Am J Health-Syst Pharm*, Vol. 58, n. 19, pp. 1824-29.
98. REASON J.T., (1990), *Human error*, Cambridge University Press, New York.
99. REASON J.T., (2000), *Human error: models and management*, *BMJ*, Vol. 320, pp. 768-70.
100. REDMILL F. et al., (1999), *Sistem Safety: HAZOP and Software HAZOP*, Ed. John Wiley & Sons.
101. RICE N., (2001), *Binomial Regression* in GOLDSTEIN H., LEYLAND A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, 27-44.
102. RICE N., LEYLAND A., (1996), *Multilevel models: applications to health data*, *Journal of Health Services Research and Policy*, 1, 154-64.
103. RICH M.W., BECKHAM V., WITTENBERG C., LEVEN C.L., FREEDLAND K.E., CARNEY R.M., (1995), *A Multidisciplinary Intervention to Prevent the Readmission of Elderly Patients with Congestive Heart Failure*, *The New England Journal of Medicine*, 333: 1190-95.
104. RICHARDS S.H., COAST J., GUNNELL D.J., PETERS T.J., POUNSFORD J., DARLOW M., (1998), *Randomised controlled trial comparing effectiveness and acceptability of an early discharge, hospital at home scheme with acute hospital care*, *British Medical Journal*, 316: 1796-1801.

105. RICHARDSON S., GREEN P.J., (1997), *On Bayesian analysis of mixtures with an unknown number of components*, Journal of the Royal Statistical Society, B 59, 731-92.
106. RIZZI L., FRANCESCUTTI C., SIMON G., BORGNOLO G., (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*, 295-310.
107. ROSENBAUM P. R., (2002), *Observational studies (2nd Edition)*, Springer-Verlag.
108. RUBIN R.J., ALTMAN W.M., MENDELSON D.N., (1994), *Health care expenditures for people with diabetes mellitus*, Journal of Clinical Endocrinology and Metabolism, 78: 809.
109. RUDD A.G., WOLFE C.D.A., TILLING K., BEECH R., (1997). *Randomised controlled trial to evaluate early discharge scheme for patients with stroke*, British Medical Journal, 315: 1039-1044.
110. SCHRAAGEN J.M. et al., (2002), *Cognitive Task Analysis*, Lawrence Erlbaum Associates, pp.1-24.
111. SCOTT R.S., BROWN L.J., AND CLIFFORD P., (1985), *Use of health services by diabetic person*, hospital admissions Diabetes Care, 8: 43-47.
112. SHEPPERD S., HARWOOD D., JENKINSON C., GRAY A., VESSEY M., MORGAN P., (1998), *Randomised controlled trial comparing hospital at home care with inpatient hospital care. I: three months follow-up of health outcomes*, British Medical Journal, 316: 1786-91.
113. SILBER J.H., ROSENBAUM P., ROSS R.N., (1995), *Comparing the Contributions of Groups of Predictors: high Outcomes Vary With Hospitals Rather Than Patient Characteristics?*, Journal of the American Statistical Association, 3 90, 7-18.
114. [SIRIO C.A.](#), [ANGUS D.C.](#), [ROSENTHAL G.E.](#), (1994), *Cleveland Health Quality Choice (CHQC), An ongoing collaborative, community-based outcomes assessment program*, New Horiz, 2 (3): 321-5.
115. SMEENK F.W.J.M., VAN HAASTREGT J.C.M., DE WITTE L.P., CREBOLDER H.F.J.M., (1998), *Effectiveness of home care programmes for patients with incurable cancer on their quality of life and time spent in hospital: systematic review*, British Medical Journal, 316: 1939-44.
116. SOLUCIENT LLC, (2003), *Solucient's 100 Top Benchmark Hospitals*, www.solucient.com.
117. STEEN P.M., BREWSTER A.C., BRADBURY R.C., ESTABROOK E., YOUNG J.A., (1993), *Predicted probabilities of hospital death as a measure of admission severity of illness*, Inquiry, 30: 128-141.

118. STEINWACHS D.M., WU A., SKINNER E.A. AND CAMPBELL D., (1995), *Managed Health Care Association Outcomes Management System Project: Asthma Patient Outcomes Study Baseline Survey Summary Report*, unpublished manuscript, January.
119. STEWART S., MARLEY J.E., HOROWITZ J.D., (1999), *Effects of multidisciplinary, home-based intervention on unplanned readmission and survival among patients with chronic congestive heart failure: a randomised controlled study*, *The Lancet*, 354: 1077-83.
120. TENNANT A., PENTA M., TESIO L., GRIMBY G., THONNARD J.L., SLADE A. et al., (2004), *Assessing and adjusting for cross-cultural validity of impairment and activity limitation scales through differential item functioning within the framework of the Rasch model: the PRO-ESOR project*, *Med Care*, 42: I37-I48.
121. TESIO L., (2003), *Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research*, *J Rehabil Med.*, 35: 105-15.
122. TESIO L., PERUCCA L., BATTAGLIA M.A, FRANCHIGNONI F., (1997), *Quality assessment of the FIM (Functional Independence Measure) Ratings Through Rash Analysis*, *Europa medicophysica*, 33, 2, 55-57.
123. THOMAS E.J. et al., (1999), *Costs of medical injuries in Utah and Colorado*, *Inquiry*, Vol. 36, n. 3, pp. 255-64.
124. THOMAS E.J. et al., (2000), *Incidence and types of adverse events and negligent care in Utah and Colorado*, *Med Care*, Vol. 38, pp. 261-71.
125. THOMAS N., LONGFORD N.T., ROLPH J.E., (1994), *Empirical Bayes methods for estimating hospital specific mortality rates*, *Statistics in Medicine*, 13, 889-903.
126. TISSOT E. et al., (1999), *Medication errors at the administration stage in an intensive care unit*, *Intensive Care Med*, Vol. 25, n. 4, pp. 353-59, April.
127. TRUCCO P., CAVALLIN M., (2003), *Clinical Risk and Error Analysis (CREA): metodo per l'analisi dei rischi organizzativi nei processi ospedalieri*, *Convegno Scientifico Nazionale Sicurezza nei Sistemi Complessi*, Bari, 16-17 Ottobre.
128. TRUCCO P., CAVALLIN M., BONINI P., RUBBOLI F., (2003), *Valutazione del rischio organizzativo nei processi di cura*, *Rischio Sanità, ASSI News*, Vol. 11, pp. 26-32, Dicembre.
129. US PHARMACOPEIA, (2000), *US Pharmacopeia summary of the 1999 information submitted to MedMARx, a national database for hospital medication error reporting*, Rockville, MD, www.usp.org
130. US PHARMACOPEIA, (2001), *US Pharmacopeia summary of the 2000 information submitted to MedMARx, a national database for hospital medication error reporting*, Rockville, MD, www.usp.org
131. VINCENT C. et al., (1998), *Framework for analysing risk and safety in clinical medicine*, *BMJ*, Vol. 316, pp. 1154-57.

132. VINCENT C. et al., (2001), *Adverse events in British hospitals: preliminary retrospective record review*, BMJ, Vol. 322, pp. 517–19, March.
133. VINCENT C., (2003), *Understanding and Responding to Adverse Events*, N engl J Med, Vol. 348, n. 11, pp. 1051-56.
134. VITTADINI G., (1999), *Analysis of qualitative variables in structural models with unique solutions*, in VICHI M., OPITZ O. (eds), *Classification and Data Analysis- Theory and Application*, Springer-Verlag, 203-210.
135. VITTADINI G., SANARICO M., (2002), *Metodologia per lo studio dell'efficacia relativa in sanità*, in *Statistica Applicata*, Vol. 14, 3, pp. 331-359.
136. VITTADINI G., ROSSI C., SANARICO M., (2003), *Recenti sviluppi nella metodologia statistica per la valutazione degli ospedali*, *Statistica*, anno LXIII, 1, in corso di pubblicazione.
137. WAGSTAFF A., VAN DOORSLAER E., (1993), *Equity in the finance and delivery of health care: concepts and definitions*, in VAN DOORSLAER E., WAGSTAFF A., RUTTEN F. (eds.), *Equity in the Finance and Delivery of Health Care: an international perspective*, Oxford University Press., 7-19.
138. WALLACE J., (1997), *Medication errors, zeroing in on medication errors [Editorial]*, *The Lancet*, Vol. 349, pp. 369.
139. WARE J.E., (1993), *Sf-36, Health Survey. Manual and Interpretation*, Guide, The Health Institute, New England Medical Center.
140. WEINBERGER M., ODDONE E.Z., HENDERSON W.G., (1996), *Does increased access to primary care reduce hospital admissions?*, *New England Journal of Medicine*, 30: 1441-7.
141. WILSON R.M. et al., (1995), *The Quality in Australian Health Care Study*, *Med J Aust*, Vol. 163, pp. 458-71.
142. WILSON A., PARKER H., WYNN A., JAGGER C., SPIERS N., JONES J., PARKER G., (1999), *Randomised controlled trial of effectiveness of Leicester hospital at home scheme compared with hospital care*, *British Medical Journal*, 319: 1542-1546.
143. WOLD H., (1982), *Soft Modeling: The Basic Design and Some Extensions*, in K.G. JÖRESKOG AND H. WOLD (eds.), *System under indirect observation: casuality, structure, prediction*, Vol. 2, 1-54.
144. WRIGHT, MASTERS, (1982), *Rating Scale Analysis*, Mesa Press, Chicago.
145. YANG M., (2001), *Multinomial Regression* in GOLDSTEIN H., LEYLAND A.H. (eds.), *Multilevel Modelling of Health Statistics*, Wiley, New York, pp. 107-126.
146. YOUNG W.W., KOHLER S., KOWALSKI J., (1994), *PMC patient severity scale: derivation and validatio*, *Health Serv Res.*, 29: 367-390.
147. ZASLAWSKY A., (2001), *Statistical issues in reporting quality data: small samples and casemix variation*, *Int J Qual Health Care*, 13 (6): 481-88.

148. ZÜRICH CONSULTING RISK MANAGEMENT, (2002), *Dossier: Quando l'errore entra in ospedale*, Milano, www.cineas.it.