

Modelli multilivello a scelta binaria: i vantaggi di una loro applicazione ai dati sanitari

Elisa Iezzi* e Cristina Ugolini**

**Dipartimento di Scienze Statistiche, Università di Bologna; **Dipartimento di Scienze Economiche, Università di Bologna e CHILD (Centre for Household Income Labour and Demographic economics)*

Riassunto. I dati sanitari frequentemente si presentano in forma multilivello, nel senso che gli individui e il loro gruppo di appartenenza possono essere considerati un sistema gerarchico. L'assunto alla base della maggior parte dei modelli statistici classici è quello di indipendenza tra le osservazioni, mentre nella realtà le osservazioni individuali inserite in ambiti gerarchici non sono del tutto indipendenti e ignorare questa dipendenza produce modelli statistici sottostimati. I modelli multilivello consentono di tenere conto della struttura gerarchica delle informazioni e di formulare, in modo adeguato, le relazioni tra il livello individuale e di gruppo. Questi modelli sono utilizzati in modo crescente a livello internazionale, mentre ancora poco diffuse sono le applicazioni a dati italiani. Scopo di questo lavoro è quello di proporre una riflessione sulle proprietà statistiche di questa metodologia con particolare enfasi al modello a scelta binaria, mettendone in luce le potenzialità di utilizzo in un contesto sanitario.

Parole chiave. Dati a struttura gerarchica, modelli logistici, modelli multilivello.

Abstract. Health outcomes research often involves analysis of data that are nested within hierarchical groups. This hierarchical data structure poses serious methodological issues that challenge the assumptions of traditional data analysis method: ignoring the dependency of clustered observations produces underspecified models. Multilevel modeling allows to examine simultaneously the effect of individual-level as well as group-level predictors on the dependent variable of interest and over the past decade has become an accepted statistical technique in health research. This paper presents an introductory study of multilevel models, focusing in particular on the multilevel logistic form, and considers their use to health data sets. To date, little use has been made of these techniques within the Italian health economics literature but the potential applications of these methods are many.

Key words. Hierarchical data structure, logistic models, multilevel modeling.

1. Introduzione

Il comportamento degli individui è spesso influenzato dal contesto sociale e le caratteristiche del gruppo a cui appartengono sono a loro volta influenzate dai singoli che compongono il gruppo stesso. Per questo motivo gli individui e il contesto sociale di appartenenza possono essere visti come un sistema gerarchico di unità e gruppi posti a livelli diversi.

I dati sanitari sono frequentemente organizzati secondo una struttura gerarchica; ad esempio, le prestazioni ospedaliere possono essere raggruppate per reparto, per ospedale, per azienda sanitaria di assi-

stenza, oppure i residenti sono iscritti alla lista di un medico di medicina generale che a sua volta è parte di un nucleo di cure primarie inserito all'interno della programmazione distrettuale e aziendale. Spesso non è semplice individuare con chiarezza il contributo dei diversi attori nel determinare un obiettivo di programmazione sanitaria, sia che si tratti di costi, di indicatori di processo o di esito clinico. Per analizzare tali contesti sono stati elaborati i modelli statistici 'multilivello' in grado di tenere conto esplicitamente della struttura gerarchica delle informazioni e di rappresentare, in modo statisticamente adeguato, le relazioni tra il livello individuale e il livello aggregato.

Questo lavoro è parte del progetto di ricerca sanitaria finalizzata ex art. 12 e 12 bis del D. Lgs. 502/92 finanziata dal Ministero della salute per l'anno 2006 dal titolo "Il Governo delle cure primarie: pagamento per risultati ed assetti organizzativi", coordinato da Francesco Taroni e svolto in collaborazione tra l'Agenzia sanitaria e sociale regionale Emilia-Romagna e i Dipartimenti di medicina e sanità pubblica e di Scienze economiche dell'Università di Bologna. Le autrici ringraziano Matteo Lippi Bruni per gli utili suggerimenti alla stesura del lavoro.

Le prime applicazioni dei modelli multilivello si trovano nelle scienze sociali, dove l'interazione tra l'individuo e l'ambiente è storicamente l'oggetto prioritario di interesse. Tale metodologia ha rappresentato il superamento e punto di incontro tra i due approcci classici opposti, ovvero quello a livello aggregato (analisi 'ecologica') e quello a livello individuale (analisi 'atomistica'), che consistono appunto nell'estendere all'individuo le relazioni emerse dai gruppi e, viceversa, nell'attribuire caratteristiche individuali a gruppi e strutture più complesse (Robinson, 1950; Diex-Roux, 2000). Lavorando a livello aggregato si elimina il problema di correlazione e non-ortogonalità delle osservazioni, ma si perde una notevole quantità di informazioni, mentre lavorando a livello individuale non si rispetta l'assunzione di indipendenza (incorrelazione ed ortogonalità) tra le osservazioni e quindi si tende a produrre stime inefficienti. Come vedremo nelle prossime pagine, le caratteristiche proprie dell'approccio multilivello consentono invece di superare i limiti derivanti, rispettivamente, dalle cosiddette fallacie ecologica e atomistica.

Un'alternativa alla metodologia multilivello potrebbe essere quella di realizzare regressioni separatamente per ogni gruppo. Questo però non consentirebbe né di valutare come le variabili a livello aggregato influenzino quelle a livello individuale, né di considerare la possibile interazione tra variabili relative a ciascun livello. Ancora, si potrebbero includere le caratteristiche di appartenenza al gruppo mediante l'utilizzo di variabili dicotomiche. Questo approccio è analogo al precedente e allo stesso modo non permette di esaminare quale sia l'influenza relativa delle diverse variabili a livello aggregato sulla variabile dipendente. Inoltre, queste soluzioni trattano i gruppi a sé stanti, non relazionati tra loro, ignorando che questi potrebbero avere caratteristiche comuni. Le analisi multilivello differiscono da questi approcci in quanto, mediante equazioni in cui sono presenti simultaneamente sia effetti a livello di gruppo che predittori a livello individuale, sono in grado di trattare adeguatamente i problemi sopracitati (Snijders e Bosker, 1994; Kelsey et al., 1996).

Scopo di questo lavoro è quello di proporre una riflessione sulla struttura e sulle proprietà dei modelli multilivello, che metta in luce i possibili vantaggi di applicazioni al contesto economico-sanitario. Data la frequente natura binaria delle variabili oggetto di indagine in tale ambito, si analizzeranno con particolare attenzione gli utilizzi della variante logistica, cercando di offrire una sintesi dei fondamenti statistici che sono alla base di questa tipologia di strumenti sempre più diffusa nell'ambito della ricerca applicata. Nell'ultima sezione del lavoro si presenterà una riflessione sistematica, seppur non esaur-

stiva, delle applicazioni proposte fino ad ora in letteratura, offrendo una rassegna ragionata dei contributi che evidenziano l'evoluzione della metodologia e le sue ancora significative potenzialità di sviluppo.

2. I modelli multilivello

I modelli multilivello derivano principalmente da due approcci: analisi di contesto (Iversen, 1991) e ad effetti misti (Longford, 1993). I modelli di contesto sono funzioni in cui la variabile dipendente è misurata a livello individuale (ad esempio, reddito personale), mentre le variabili esplicative possono essere sia individuali (ad esempio, grado di istruzione), che variabili di gruppo, chiamate anche variabili di contesto (ad esempio, variabili ambientali o geografiche). Questi modelli sono considerati ad effetti fissi poiché non considerano l'esistenza di una variabilità a livello di gruppo, ma inseriscono le variabili di contesto come media del gruppo stesso. In questo modo la variabilità tra gruppi si comprime e può causare problemi di multicollinearità tra le variabili. I modelli ad effetti misti possono considerarsi un successivo sviluppo dei modelli di contesto in grado di riconoscere la dipendenza tra le osservazioni appartenenti ad uno stesso gruppo e di risolvere i problemi di multicollinearità. Tuttavia, i residui di questa tipologia di modelli continuano ad essere definiti a livello individuale e quindi sono indipendenti e distribuiti come una normale a media nulla e varianza costante. Rispetto a questi due approcci, i modelli multilivello compiono un passo avanti poiché inseriscono la variabilità a livello gruppo, permettendo così un legame degli individui all'interno dei gruppi.

La problematica statistica principale trattata mediante stime multilivello si origina ogni qualvolta le variabili esplicative sono riferibili a unità di osservazione raggruppabili in aggregati di livello superiore. Ad esempio, negli studi di carattere economico-sanitario accade frequentemente che, oltre a informazioni relative a singoli soggetti (dati individuali), si disponga di dati relativi a specifiche forme organizzative ovvero ad aggregati territoriali. In questi casi, si genera una potenziale correlazione fra le osservazioni di cui la metodologia consente di tenere conto. Ad esempio, le caratteristiche (non osservabili) di pazienti appartenenti al medesimo medico presentano presumibilmente una maggiore correlazione rispetto a pazienti che fanno capo a medici diversi. Analogamente, medici operanti nel medesimo ambito territoriale possono presentare livelli di omogeneità per una varietà di caratteristiche (non osservabili) superiori rispetto a quei medici operanti in territori diversi. Le tradizionali tecniche di stima finiscono per non tenere conto del maggiore grado di correlazione delle osservazioni che fanno riferimento alla medesima

unità di livello superiore. Infatti, l'assunto alla base di molti modelli statistici classici è quello di indipendenza tra le osservazioni, mentre, nella realtà, le osservazioni individuali inserite in ambiti gerarchici non sono del tutto indipendenti. Violando questa ipotesi, i modelli classici non possono essere utilizzati poiché produrrebbero una sottostima della variabilità e quindi una significatività del test del tutto inappropriata. Invece, la metodologia multilivello consente di individuare quanta parte della variabilità totale è da attribuire ai fattori di ciascun livello, e anche l'esistenza di relazioni diverse tra le caratteristiche individuali e di contesto (cioè attribuibili all'appartenenza a un dato gruppo) e le loro reciproche relazioni (interazione tra livelli). In questo modo è possibile studiare in che modo l'aggregato 'ambiente' condiziona direttamente o indirettamente le scelte, i comportamenti, i processi decisionali dell'individuo.

La scelta dei livelli della gerarchia, e quindi degli ambiti di influenza sul comportamento individuale, è un aspetto cruciale sia dal punto di vista sostanziale per l'interpretazione dell'effetto dei gruppi, sia dal punto di vista statistico in quanto si corre il rischio di sovrastimare l'impatto dell'aggregato considerato (Goldstein, 1995). Tuttavia, in ambito sanitario le macro-situazioni tendono a costituire già di per sé stesse un raggruppamento e quindi una gerarchia 'naturale', che rende abbastanza semplice l'individuazione dei diversi livelli. Ad esempio, se si considera il caso dei pazienti di una regione, questi possono essere visti come un'organizzazione articolata su quattro livelli. La struttura può essere graficamente rappresentata come visualizzato nella figura 1.

I pazienti hanno un unico medico di riferimento, ogni medico partecipa all'attività di programmazione sanitaria attraverso il proprio distretto sanitario, che è articolazione organizzativa di un'azienda territoriale. I pazienti sono quindi il primo livello (*micro*) di una struttura gerarchica che l'analisi multilivello è in grado di riflettere, mentre medici,

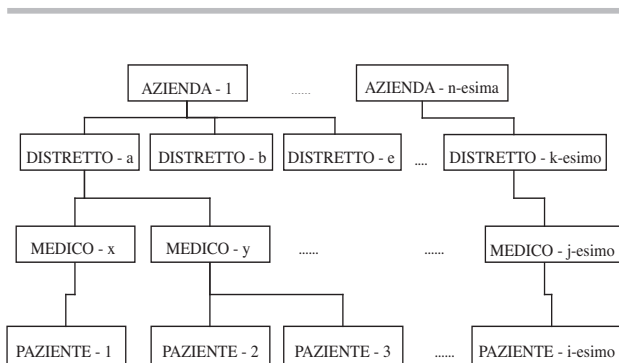


Figura 1 - Rappresentazione schematica di una struttura gerarchica.

distretti e aziende rappresentano le unità di livello superiore (*macro*).

2.1. I modelli di regressione multilivello

Il modello di regressione multilivello è conosciuto in letteratura con nomi diversi, ad esempio *random coefficient model*, *variance component model* e *hierarchical linear model*. Tutti questi modelli assumono una struttura gerarchica dei dati con una variabile dipendente misurata a livello individuale e variabili esplicative su tutti i livelli. Concettualmente il modello è spesso visto come un sistema di regressioni gerarchiche.

Ipotizzando di rappresentare una semplice gerarchia su due livelli, la struttura di base del modello multilivello è la seguente:

$$y_{ij} = g(\beta'x_{ij}) + u_j + e_{ij}$$

dove y_{ij} rappresenta la variabile dipendente, che è messa in relazione con le variabili esplicative attraverso la funzione $g(\cdot)$, che esprime appunto la forma funzionale del legame tra variabile dipendente e variabili esplicative. Gli errori casuali hanno media zero e varianza costante $cov(e_{ij}, u_j) = 0$ e cioè i livelli sono incorrelati. Per l' i -esimo micro-livello annidato entro il j -esimo macro-livello il modello avrà come valore atteso $E(y_{ij} | x_{ij}) = \beta'x_{ij}$ e come varianza $var(y_{ij} | x_{ij}) = \sigma_e^2 + \sigma_u^2$ condizionata. La varianza totale sarà ripartita in due componenti e la struttura della matrice di varianza e covarianza sarà una matrice a blocchi:

$$cov(y_{ij} | \mathbf{x}) = \begin{pmatrix} \mathbf{A}_1 & \dots & \mathbf{0} \\ \dots & \dots & \dots \\ \mathbf{0} & \dots & \mathbf{A}_M \end{pmatrix}$$

dove

$$\mathbf{A}_j = \begin{pmatrix} \sigma_e^2 + \sigma_u^2 & \dots & \sigma_u^2 \\ \dots & \dots & \dots \\ \sigma_u^2 & \dots & \sigma_e^2 + \sigma_u^2 \end{pmatrix} \quad j = 1 \dots M$$

La partizione della varianza offre la possibilità di calcolare la correlazione *intracluster*

$$\frac{\sigma_u^2}{\sigma_u^2 + \sigma_e^2}, \text{ dato che}$$

$$cov(u_j + e_{i1j}, u_j + e_{i2j}) = cov(u_j, u_j) = \sigma_u^2.$$

La correlazione *intracluster* (ρ), o coefficiente di *Intra-Class Correlation* (ICC), misura il legame tra due soggetti dentro lo stesso gruppo e quindi rappresenta la forza del legame (*nesting*) dei dati gerarchici. Quando il coefficiente è elevato, allora si può affermare l'esistenza di una variabilità del gruppo si-

gnificativa che non potrebbe essere spiegata qualora non si considerassero le aggregazioni per livelli. Lo stimatore generalmente utilizzato è quello di massima verosimiglianza. I residui sono assunti con distribuzione normale a media zero. Essendo generati da una procedura di massima verosimiglianza hanno bisogno di ampi campioni, in tutti i livelli, per evitare una loro sottostima (Maas e Hox, 2005). In presenza di gruppi poco numerosi è possibile adottare la strategia di trasferire le informazioni dai gruppi più numerosi a quelli meno numerosi, in modo tale da consentire comunque l'applicazione della procedura di massima verosimiglianza (Snijders e Bosker, 1993).

La principale caratteristica dell'approccio multilivello è quella di introdurre correzioni nella stima della matrice varianza/covarianza degli errori in grado di assicurare un trattamento appropriato dei dati quando le variabili esplicative sono annidate le une nelle altre. Rispetto ai tradizionali modelli di regressione tale approccio assicura quindi una serie di vantaggi nel processo inferenziale (Guo e Zhao, 2000). Oltre ad introdurre una correzione dei coefficienti, permette anche di ottenere errori standard che tengono conto dei diversi gradi di correlazione ai differenti livelli e di conseguenza assicura il calcolo di corretti intervalli di confidenza e test di significatività. In altre parole, l'analisi multilivello permette di valutare l'influenza di ciascuna unità di livello superiore sullo stato della variabile dipendente dalle unità di primo livello (Aitkin e Longford, 1986), mentre quantificare la variabilità dei macro errori fornisce la possibilità di avere variabili dipendenti a livello individuale che potrebbero, dentro lo stesso gruppo, essere correlate (Goldstein, 1995).

Considerare l'idea di gruppi ordinati secondo una struttura gerarchica consente anche di migliorare il *design effect* (*deff*) (Kish, 1965; Balderjahn et al., 1998), che spiega l'efficacia in termini di varianza dell'utilizzo dei gruppi rispetto al non considerare livelli macro. In un semplice campionamento a gruppi con uguale dimensione dei gruppi il *deff* può essere calcolato come: $deff = [1 + \rho(n_{clus} - 1)]$ dove ρ è il coefficiente *intracluster* e n_{clus} è la dimensione dei gruppi. È chiaro che $deff = 1$ solo se non c'è correlazione dentro i gruppi, cioè $\rho = 0$ o la dimensione del cluster sia nulla, $n_{clus} = 0$. In tutte le altre situazioni il *deff* è maggiore di uno, il che implica che le formule statistiche standard sottostimano la varianza campionaria quando si considerano gerarchie e quindi aggregazioni (gruppi) di macro livello.

Dopo aver valutato l'opportunità di applicare ai dati un'analisi multilivello, si dovrà decidere quale sia la forma della funzione che lega la variabile dipendente con i residui. La prima questione da affrontare è quale sia il tipo di variabilità da attribuire all'intercetta e ai coefficienti delle covariate. Se si ha

ragione di ipotizzare che la variabilità debba essere relativa unicamente all'intercetta, la funzione sarà espressa da un insieme di rette di regressione (una per ogni gruppo) tra loro parallele, a cui è associato un ordinamento (punto di origine della retta) tra i gruppi, dettato esclusivamente dal termine u_j . Il modello diventa più complesso se la componente casuale viene attribuita anche al coefficiente β ('modello a pendenza variabile'). In pratica se si considerasse una funzione lineare si potrebbe immaginare di avere un fascio di rette con la stessa origine a pendenza variabile. Se invece si considerano variabili sia l'intercetta che i coefficienti, allora vi sarà un numero maggiore di termini nel modello e le rette di regressione presenteranno origini e inclinazioni differenti tra loro (in quanto diversi possono essere i coefficienti β). In questo caso, si suppone un effetto variabile dei gruppi di variabili esplicative su quella dipendente ed anche il valore medio della variabile dipendente cambierà in base al gruppo di appartenenza. Questi modelli sono chiamati 'a coefficienti variabili' e sono un'estensione dei primi due, fascio di rette a pendenza variabile che si formano in punti ad altezze diverse.

Nella scelta della forma funzionale, dopo aver attribuito il tipo di variabilità si dovranno analizzare i dati a cui applicare il modello e quindi scegliere il legame tra variabile dipendente e variabili esplicative. Questa scelta dipenderà dal tipo di variabile dipendente, se continua, discreta binaria o a scelta multipla, ma anche dalle osservazioni a disposizione, se di tipo *cross-sectional* (DiPrete e Grusky, 1990) o *time-series* (Gibbons et al., 1988; Goldstein et al., 1994). Infatti, quando gli stessi individui sono misurati in più occasioni allora si potrebbero adottare modelli multilivello in cui gli individui rappresentano il livello aggregato e le diverse rilevazioni il livello individuale. Quando le osservazioni sono ripetute nel tempo, allora ogni individuo presenta una personale traiettoria, e quindi la variabile tempo ricade fra le unità di primo livello. Un altro tipo di applicazione si ha quando la variabile dipendente è discreta. In questo caso cambia la forma del modello che non sarà più lineare come nel caso continuo, ma di tipo non lineare come logit/probit se la variabile discreta è a scelta binaria, oppure modello di poisson multilivello se la variabile è di 'count'. Quando la variabile dipendente è rappresentata da più di due categorie avremo un caso di scelta multipla e ricadremo all'interno dei modelli multinomiali multilivello (Goldstein e Rasbash, 1992).

2.2. Il modello logistico multilivello

In tutti i modelli multilivello lineari si assume che la variabile risposta abbia distribuzione continua e i coefficienti casuali e i residui siano normalmente

distribuiti. Quando la variabile dipendente è una variabile a scelta binaria (dicotomica) allora il modello preferito è quello logistico in cui $\text{logit}(p)$ risulta una funzione lineare delle variabili esplicative.

Nei modelli binari la parte casuale individuale è di tipo logistico ed invece la parte casuale del gruppo ha distribuzione normale (Duncan et al., 1998). Se si considera una struttura gerarchica organizzata su tre livelli dove con l'indice i si indicano i pazienti (livello 1), con j si indicano gli ospedali (livello 2) e con k le regioni (livello 3), la y_{ijk} sarà la variabile dipendente di un modello multilivello binario. Tale variabile, che ha probabilità π_{ijk} per l'esito 1 e $1-\pi_{ijk}$ per l'esito 0, segue una distribuzione binomiale con media e varianza nota.

La probabilità di risposta può essere modellata utilizzando il modello *logit*, tale che:

$$\log \left[\frac{\pi_{ijk}}{1-\pi_{ijk}} \right] = \beta_0 + \beta_1 \mathbf{x}_{ijk} + v_{0k} + u_{0jk} + e_{ijk}$$

Condizionando rispetto u_{0jk} e rispetto v_{0k} , le x_{ijk} sono variabili esplicative. La variabile π_{ijk} è la probabilità che il paziente i sia ricoverato nell'ospedale j della regione k . In questo modello la parte casuale è quella degli errori $v_{0k} + u_{0jk} + e_{ijk}$ che rappresentano rispettivamente termini di errori casuali a livello 3, 2, 1. In generale si assume che queste quantità casuali siano incorrelate e con distribuzione normale $v_{0k} \sim N(0, \sigma_v^2)$ e $u_{0jk} \sim N(0, \sigma_u^2)$. Invece la varianza del primo livello assume il valore fisso

$$\sigma_e^2 = \frac{\pi^2}{3}$$

che indica la varianza della distribuzione logistica standard (Guo e Zhao, 2000). Le stime σ_v^2 e σ_u^2 del modello indicano il grado di variabilità nei due livelli, regioni e ospedali. Se una stima di queste varianze non risulta significativa allora è possibile escludere dall'analisi il relativo livello.

Per valutare l'importanza della variabilità all'interno di ogni livello, Browne et al. (2005) rinominano il coefficiente *intracluster* come *Variance Partition Coefficient*:

$$VPC_{ijk} = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_u^2 + \frac{\pi^2}{3}}$$

varianza tra i livelli k e

$$VPC_{ijk} = \frac{\sigma_u^2}{\sigma_v^2 + \sigma_u^2 + \frac{\pi^2}{3}}$$

varianza tra i livelli j , interpretabili come la frazione di variabilità totale attribuibile a quel determinato livello.

Per stimare i parametri del modello logistico i metodi di stima utilizzati per i modelli lineari non

possono essere direttamente applicati, ma è necessario rendere lineare la funzione esponenziale per poter poi applicare una stima di quasi-verosimiglianza (McCullagh e Nelder, 1989). In generale, vengono utilizzati metodi che approssimano la verosimiglianza come *Marginal Quasi-Likelihood* (Goldstein e Rasbash, 1996), *Penalized Quasi-Likelihood* (Breslow e Clayton, 1993) e, più recentemente, metodi che utilizzano la quadratura gaussiana (Pinheiro e Bates, 1995) in grado di ottenere stime accurate anche per piccoli campioni (Paccagnella, 2006).

Una volta stimati i parametri, la loro interpretazione risulterà particolarmente impegnativa a causa della non-linearità del modello logistico. Il metodo classico per l'interpretazione dei coefficienti nei modelli logistici è l'*odds ratio* (Snijders e Bosker, 1999), definito come il rapporto tra la probabilità di successo sulla probabilità di insuccesso. Una delle trasformazioni delle probabilità maggiormente utilizzate per calcolare un *odds ratio* è definita da:

$$\text{logit}(p) = \ln \left(\frac{p}{1-p} \right)$$

dove $\ln(x)$ è il logaritmo naturale di x . Il vantaggio dell'*odds ratio* è quello di lavorare su un'altra unità di scala che può assumere valori non limitati ad un intervallo 0-1. In questo modo i valori stimati dei parametri esprimono gli effetti esercitati sul *logit* rispettivamente delle singole variabili, delle loro possibili iterazioni e dei valori dei residui.

Nei modelli multilivello ogni residuo è ipotizzato essere variabile casuale avente una distribuzione con parametro fisso che individua la variabilità nel livello. Per ciascun livello gerarchico, i valori dei residui e di quelli standardizzati rispetto ai corrispondenti errori standard possono essere utilizzati, impiegando opportune rappresentazioni grafiche, per verificare le assunzioni poste nel modello, per individuare le singole unità di un determinato livello che presentano residui molto elevati o molto bassi, e per svolgere un'analisi basata sul confronto tra coppie di unità dello stesso livello che consente di individuare gruppi di unità simili o differenti (Goldstein e Spiegelhalter, 1996).

Anche per questi modelli la numerosità dei gruppi deve essere generalmente alta. Moineddin et al. (2007) mostrano che le stime non sono affette da errore se le dimensioni dei gruppi sono di almeno 50 unità. In questo modo si riuscirà a non sottostimare le covarianze e nemmeno gli errori standard.

3. Le applicazioni in economia sanitaria

Rice e Jones (1997) illustrarono i vantaggi di una diffusione estesa dell'approccio multilivello all'am-

bito sanitario. Negli anni successivi le applicazioni di questa metodologia nella ricerca empirica sui servizi sanitari sono state molteplici, anche se forse non ancora così numerose come in altri ambiti delle scienze sociali. Utilizzando l'area di applicazione come criterio di classificazione dei lavori, vediamo brevemente i temi principali e alcuni riferimenti bibliografici. Data la significativa numerosità delle applicazioni multilivello in alcune aree, generalmente abbiamo preferito citare i primi lavori sviluppati nella seconda metà degli anni '90, alcuni di questi citati anche da Rice e Jones, e dare conto poi degli sviluppi più recenti che si concentrano su temi più specifici e/o su singole patologie e ampliano la ricerca anche a questioni in precedenza trascurate come ad esempio l'area della prevenzione. Dal punto di vista statistico, lo sviluppo di nuovi pacchetti informatici, sempre più di facile utilizzo (ad esempio, software dedicati alla metodologia multilivello come MIWin, HLM, oppure applicazioni multilivello di software generali come SAS, SPSS e STATA con il modulo gllamm per variabili latenti), ha permesso un maggiore sfruttamento delle potenzialità di calcolo di questi programmi anche a ricercatori meno esperti.

Una delle principali aree di applicazione delle tecniche multilivello è rappresentata dal tema della variabilità prescrittiva (*medical variation*), secondo il quale le probabilità che un paziente affetto da una determinata malattia riceva il medesimo trattamento recandosi in strutture diverse o rivolgendosi a medici diversi sono molto basse e le variazioni di utilizzo dei servizi sanitari sono spesso molto grandi anche entro aree geografiche di limitata estensione. Da ciò scaturisce l'ipotesi che le cause dell'ineguale distribuzione delle risorse non sia una questione legata all'incidenza delle malattie, bensì un fenomeno legato alla pratica individuale dei medici. L'obiettivo di ridurre la variabilità delle cure, almeno per le patologie più frequenti, impone di comprendere meglio le cause che sono alla base di questa variabilità. Da questo punto di vista, le potenzialità offerte dall'analisi multilivello appaiono particolarmente appropriate in relazione agli specifici obiettivi conoscitivi e portano con sé l'auspicio di un crescente utilizzo dello strumento in quest'area della ricerca economica ed epidemiologica. Tra i diversi lavori che hanno analizzato la variabilità prescrittiva utilizzando questa metodologia sul tema, si può citare quello di Davis e Gribben (1995) che studia, su dati neozelandesi, la variabilità delle prescrizioni farmaceutiche. L'analisi è stata poi successivamente estesa anche ad altre possibili fonti di variabilità riconducibili ai medici di medicina generale (esami di laboratorio, radiografie, invito ad una visita di follow-up, etc.) (Davis et al., 2002). Nello stesso periodo due lavori fra loro collegati hanno analizzato su dati australiani

l'influenza sulla variabilità prescrittiva di diversi meccanismi di remunerazione dei medici di medicina generale (Scott e Shiell, 1997a) e della presenza di diversi gradi di concorrenza territoriale (Scott e Shiell, 1997b). Su dati statunitensi, lo studio recente di Webster et al. (2008) cerca di valutare la variabilità nella prescrizione di oppiacei per dolori alla schiena di varia intensità o i lavori di Verstappen et al. (2005) e di De Bakker et al. (2007) sulla variabilità prescrittiva dei medici olandesi. Per la Francia, il lavoro di Pelletier et al. (2007) indaga le determinanti in grado di influenzare le vaccinazioni eseguite dai medici in ambulatorio. Molteplici sono poi le applicazioni a singole patologie, ad esempio le indagini sulla variabilità prescrittiva relativamente ai pazienti diabetici (Krein et al., 2002) o ipertesi (Johnell et al., 2005). Numerose indagini riguardano l'assistenza secondaria di tipo specialistico [ad esempio, Luo et al. (2007) esaminano le diverse propensioni dei chirurghi a inviare i propri pazienti ai medici oncologi in presenza di tumore] o ospedaliera, come il diverso grado di diffusione di una certa pratica clinica [ad esempio, il cesareo (Brown, 2007), la laparoscopia (Hollingworth et al., 2008) o lo stent medicato (Austin et al., 2008)].

Parallelamente al tema della variabilità prescrittiva, i lavori di Von Korff et al. (1992) e Duncan et al. (1996) dimostrano come il comportamento adottato da un individuo rispetto alle proprie scelte in materia di salute e stile di vita si modifichi in relazione al contesto geografico in cui l'individuo vive. Di questo fatto è necessario tenere conto quando si esaminano le scelte di salute individuali e le loro conseguenze sugli esiti sanitari. Questa letteratura dimostra come le scelte di salute siano influenzate dal luogo di provenienza e/o di residenza e come le interazioni tra le caratteristiche ambientali e le caratteristiche individuali siano cruciali per comprendere meglio gli atteggiamenti rispetto, ad esempio, al consumo di alcol tra gli adolescenti (Demers et al., 2002 per il Canada; Kuntsche et al., 2008 per la Svizzera), al consumo di tabacco o droghe (Williams e Latkin, 2007), al ricorso a metodi contraccettivi [dopo il lavoro pionieristico di Entwisle et al. (1986), negli anni successivi numerosi lavori hanno esaminato le scelte contraccettive delle donne soprattutto nei paesi in via di sviluppo, ad esempio Kaggwa et al. (2008)]. Tra gli studi più recenti appartenenti a questo filone risultano di particolare interesse il lavoro di Bernburg et al. (2008), sui fattori sociali che possono influenzare le decisioni di suicidio tra gli adolescenti, e quello di Kogan et al. (2008) sulle variabili che influenzano le scelte di allattamento materno.

Si riscontrano anche forti analogie con una serie di lavori che cercano di stimare come il contesto am-

bientale influenzi la propria percezione dello stato di salute individuale. Un esempio in questo campo è l'indagine di Iversen (2008), che si pone come obiettivo quello di stimare l'associazione tra il capitale sociale (ad esempio, il tasso di partecipazione al voto o alla vita religiosa della propria comunità) e lo stato di salute percepito in Norvegia. Su questioni non dissimili, numerosi lavori studiano la relazione tra disuguaglianza del reddito e stato di salute, evidenziando come i paesi caratterizzati da maggiori disuguaglianze socioeconomiche presentino rischi di mortalità più elevati per tutte le cause di morte (Lochner et al., 2001), una peggiore autovalutazione del proprio stato di salute (Lopez, 2004) e una maggiore prevalenza di sintomi depressivi (Kahn et al., 2000). Recentemente, si è passati ad esaminare se gli effetti di una diversa distribuzione del reddito tra aree geografiche sullo stato di salute sia più pronunciato per certi gruppi di popolazione piuttosto che altri (Subramanian e Kawachi, 2006).

Un'area ulteriore di frequente applicazione dei modelli multilivello è legata alle possibilità di miglioramento delle formule allocative per la redistribuzione delle risorse sanitarie dal centro alla periferia. Queste formule sono basate prevalentemente sull'utilizzo di indici di deprivazione utilizzati come *proxies* del bisogno di salute di una certa area geografica. L'ipotesi assume che la relazione tra bisogni di salute e deprivazione sia costante tra le aree, mentre è possibile in realtà che esistano ulteriori fonti di variabilità del bisogno sanitario che non sono esaurientemente spiegate dagli indici di deprivazione. Attraverso un modello multilivello, Congdon (1995) analizza una serie di distretti elettorali inglesi appartenenti a 53 distretti amministrativi a Londra e nell'Inghilterra orientale. L'analisi dimostra l'esistenza di una significativa variabilità non spiegata tra le diverse aree, suggerendo la necessità di studiare meccanismi allocativi più complessi, in grado di tenere maggiormente conto delle diverse fonti di variabilità tra le aree. Tra i lavori più recenti in questo filone, si possono citare Barnett et al. (2001), Sundquist et al. (2004) e Congdon (2009) in cui, al fine di migliorare i criteri allocativi utilizzabili nella sanità pubblica, viene proposto di calcolare con la metodologia multilivello indicatori di prevalenza di diverse patologie croniche nel contesto statunitense.

Si registrano anche importanti applicazioni alla valutazione di efficienza ed efficacia (*performance*) dell'assistenza sanitaria nonché alla standardizzazione del rischio (*risk adjustment*) per la costruzione delle cosiddette 'league tables', che classificano strutture e interventi. Relativamente al primo tema, una delle prime applicazioni del modello multilivello è costituita dal lavoro di Goldstein e Spiegelhalter (1996), che studia l'efficacia di differenti chirurgici

in interventi di bypass aortocoronarico. Altri esempi di applicazione, il lavoro di Duncan et al. (1998) che mostra 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; il lavoro di Leyland e Boddy (1998), che studia 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 comorbidità. Sulla stessa linea, ma più recente, la valutazione dell'efficacia di 21 sistemi sanitari effettuata da Or et al. (2005), utilizzando ancora come variabile dipendente il tasso di mortalità per scompenso cardiaco. Relativamente alla standardizzazione del rischio necessaria per la costruzione delle classifiche, Austin et al. (2002) valutano il grado di concordanza tra graduatorie di ospedali costruite sulla base di quattro diversi metodi di valutazione ottenuti attraverso diverse tecniche di stima, tra cui la metodologia multilivello, applicati a un campione di pazienti ricoverati per infarto acuto al miocardio in 152 ospedali canadesi.

Numerose applicazioni hanno riguardato l'ambito ospedaliero o l'impatto prodotto da diversi modelli organizzativi. Leung et al. (1998) si propongono di identificare i fattori in grado di influenzare la durata della degenza post-parto su dati californiani, mentre Cho (2003) costruisce un modello a due livelli per testare l'ipotesi che una maggiore disponibilità di personale infermieristico riduca la probabilità che i pazienti si ammalino di polmonite durante il ricovero ospedaliero. Tra i lavori più recenti, Giorda et al. (2006) valutano se l'utilizzo delle cure specialistiche nei centri diabetici abbia effetto sulla frequenza di ospedalizzazioni per diabete, evidenziando come l'intensità dell'assistenza specialistica influenzi le probabilità di un'ospedalizzazione. Relativamente al contesto statunitense, molteplici studi hanno inoltre applicato la metodologia multilivello per verificare l'impatto prodotto dalla modificazione dei sistemi di regolamentazione o di remunerazione adottati nell'ambito dei programmi pubblici sull'assistenza ospedaliera, sulle residenze per anziani o sull'assistenza sanitaria erogata ai minori (Shenkman et al., 2003; Intrator et al., 2007).

Infine, recentemente la metodologia multilivello è stata applicata anche al tema della valutazione economica in sanità che fornisce indicazioni sulla capacità di interventi alternativi di incrementare il benessere sociale. Una discussione approfondita delle possibilità di migliorare le analisi costo-efficacia si trova, ad esempio, in Manca et al. (2005) e in Grieve et al. (2007).

4. Conclusioni

L'analisi multilivello consente di tenere conto dell'esistenza di gruppi gerarchici formati da diverse unità statistiche, di stimare la variabilità dei livelli della gerarchia e gli effetti dei gruppi sulla variabile risposta. Infine, assicura l'accuratezza delle stime sia per gli errori standard che per i parametri. Per questi motivi, i modelli multilivello risultano estremamente attraenti e trovano già oggi diffusa applicazione in ambito sanitario, dove spesso i dati sono strutturati in modo gerarchico e il fenomeno oggetto di indagine non dipende solo da variabili individuali, ma anche da relazioni ambientali di diverso genere. Della letteratura internazionale di carattere economico-sanitario che propone applicazioni multilivello abbiamo parlato senza pretese di esaustività nella sezione precedente, mentre molto più scarse sono le sue applicazioni al contesto italiano. Per alcuni esempi molto recenti si può vedere il lavoro di Agabiti et al. (2008), che misura il grado di associazione tra la condizione socioeconomica dei pazienti e le complicazioni post-operatorie in seguito ad alcune tipologie di interventi cardiovascolari, esaminando le schede di dimissione ospedaliera di 19.310 pazienti trattati in quattro città italiane nel periodo 1997-2000; il lavoro di Lippi Bruni et al. (2009), che valuta l'efficacia di diversi programmi di incentivo economico nell'ambito della medicina generale, disponendo di dati della regione Emilia-Romagna per l'anno 2003; il lavoro di Monzani et al. (2008), che studia l'efficacia dei servizi mentali in Lombardia, analizzando i percorsi di cura di 4712 pazienti trattati nell'anno, 2000 in 10 Dipartimenti di salute mentale.

Un maggiore utilizzo della metodologia multilivello sarebbe particolarmente auspicabile per il nostro paese. Infatti, soprattutto nei contesti istituzionali orientati alla programmazione, la sua applicazione appare particolarmente promettente per ricavare interessanti indicazioni di *policy* sul ruolo di alcune recenti aggregazioni istituzionali (ad esempio, le diverse forme di associazionismo medico, i nuclei delle cure primarie a livello distrettuale, le reti ospedaliere organizzate secondo il modello hub & spoke, le organizzazioni sovra-aziendali del tipo aree vaste, etc.). Al tempo stesso, è necessario lavorare nella direzione di sviluppare teorie capaci di individuare a livello di gruppo e a livello individuale i fattori che congiuntamente contribuiscono alla determinazione di un risultato; quindi la ricerca empirica dovrebbe svilupparsi di pari passo con una maggiore conoscenza dei processi alla base di un determinato fenomeno.

Va comunque ammesso che, come tutti i modelli statistici, anche i modelli multilivello necessariamente semplificano processi complessi e per questo motivo presentano dei limiti. Ad esempio, un limite intrinseco

che l'analisi multilivello condivide con gli altri metodi di regressione è il fatto di verificare separatamente per gruppo di appartenenza gli effetti delle variabili esplicative sulla variabile dipendente. Ciononostante, la metodologia multilivello continua ad essere uno dei metodi statistici più appropriati per trattare un'ampia gamma di problemi tipicamente osservabili nei dati sanitari e, per questo motivo, la ricerca in questo ambito necessita di sforzi ulteriori nella direzione di migliorare sempre più la comprensione delle dinamiche comportamentali in gioco e delle loro diverse interazioni.

Bibliografia

- Agabiti N, Cesaroni G, Picciotto S et al. (2008), The association of socioeconomic disadvantage with postoperative complications after major elective cardiovascular surgery, *J Epidemiol Community Health*, 2 (10): 882-889.
- Aitkin M, Longford N (1986), Statistical modelling issues in school effectiveness studies, *JRSS Ser A*, 149: 1-43.
- Austin D, Oldroyd KG, McConnachie A et al. (2008), Hospital and operator variations in drug-eluting stent use: a multi-level analysis of 5967 consecutive patients in Scotland, *J Public Health*, 30 (2): 186-193.
- Austin PC (2002), A comparison of bayesian methods for profiling hospital performance, *Med Decis Making*, 22: 163-172.
- Balderjahn I, Mathar R, Schader M (1998), *Classification, data analysis, and data highways*, New York, Springer Verlag.
- Barnett S, Roderick P, Martin D, Diamond I (2001), A multilevel analysis of the effects of rurality and social deprivation on premature limiting long term illness, *J Epidemiol Community Health*, 55 (1): 44-51.
- Bernburg JG, Thorlindsson T, Sigfusdottir ID (2008), The spreading of suicidal behavior: the contextual effect of community household poverty on adolescent suicidal behavior and the mediating role of suicide suggestion, *Soc Sci Med*, 68 (2): 380-389.
- Breslow N, Clayton D (1993), Approximate inference in generalized linear mixed models, *J Am Stat Assoc*, 88: 9-25.
- Brown HS 3rd (2007), Lawsuit activity, defensive medicine, and small area variation: the case of Cesarean sections revisited, *Health Econ Polic Law*, 2: 285-296.
- Browne W, Subramanian S, Jones K, Goldstein H (2005), Variance partitioning in multilevel logistic models, *JRSS Ser A*, Pt3: 599-613.
- Cho S (2003), Using multilevel analysis in patient and organizational outcomes research, *Nurs Res*, 52 (1): 61-65.
- Congdon P (1995), The impact of area context on long term illness and premature mortality: an illustration of multi-level analysis, *Reg Stud*, 29: 327-344.
- Congdon P (2009), A multilevel model for cardiovascular disease prevalence in the US and its application to micro area prevalence estimates, *Int J Health Geographics*, 8: 6.
- Davis P, Gribben B (1995), Rational prescribing and interpractitioner variation. A multilevel approach, *Int J Technol Assess*, 11 (3): 428-442.
- Davis P, Gribben B, Lay-Yee R, Scott A (2002), How much variation in clinical activity is there between general practitioners? A multi-level analysis of decision-making in primary care, *J Health Serv Res Policy*, 7 (4): 202-208.
- De Bakker DH, Coffie DS, Heerdink ER et al. (2007), Determinants of the range of drugs prescribed in general practice: a cross-sectional analysis, *BMC Health Serv Res*, 7: 132.
- Demers A, Kairouz S, Adlaf EM et al. (2002), Multilevel analy-

- sis of situational drinking among Canadian undergraduates, *Soc Sci Med*, 55 (3): 415-424.
- Diex-Roux A (2000), Multilevel analysis in public health research, *Ann Rev Publ Health*, 21: 171-192.
- DiPrete T, Grusky D (1990), The multilevel analysis of trends with repeated cross-sectional data, *Socio Methodol*, 20: 337-368.
- Duncan C, Jones K, Moon G (1996), Health-related behavior in context: a multilevel modeling approach, *Soc Sci Med*, 42: 817-830.
- Duncan C, Jones K, Moon G (1998), Context, composition, and heterogeneity: using multilevel models in health research, *Soc Sci Med*, 46: 97-177.
- Entwisle B, Mason W, Hermalin H (1986), The multilevel dependence of contraceptive use on socioeconomic development and family planning program strength, *Demography*, 23: 199-216.
- Gibbons R, Hedeker D, Watermaux C (1988), Random regression models: a comprehensive approach to the analysis of longitudinal psychiatric data, *Psychopharmacol Bull*, 24: 438-443.
- Giorda C, Petrelli A, Gnani R (2006), The impact of second-level specialized care on hospitalization in persons with diabetes: a multilevel population-based study, *Diabetic Med*, 23: 377-383.
- Goldstein H (1995), *Multilevel statistical models* (Second edition), Kendall's Library of Statistics 3, Londra, Edward Arnold.
- Goldstein H, Healy M, Rasbash J (1994), Multilevel time series models with applications to repeated measures data, *Stat Med*, 13: 1643-1655.
- Goldstein H, Rasbash J (1992), Efficient computational procedures for the estimation of parameters in multilevel models based on iterative generalised least squares, *Comput Stat Data An*, 13: 63-71.
- Goldstein H, Rasbash J (1996), Improved approximations for multilevel models with binary responses, *JRSS Ser A*, 159: 505-513.
- Goldstein H, Spiegelhalter DJ (1996), League tables and their limitations: statistical issues in comparison of institutional performance, *JRSS Ser A*, 159 (3): 385-443.
- Grieve R, Nixon R, Thompson SG, Cairns J (2007), Multilevel models for estimating incremental net benefits in multinational studies, *Health Econ*, 16: 815-826.
- Guo G, Zhao H (2000), Multilevel modeling for binary data, *Annu Rev Sociol*, 26: 441-462.
- Hollingsworth JM, Krein SL, Dunn RL et al. (2008), Understanding variation in the adoption of a new technology in surgery, *Med Care*, 46 (4): 366-371.
- Intrator O, Grabowski DC, Zinn J et al. (2007), Hospitalization of nursing home residents: the effects of states' Medicaid payment and bed-hold policies, *Health Serv Res*, 42 (4): 1651-1671.
- Iversen G (1991), *Contextual analysis*, Newbury Park, Sage University Press.
- Iversen T (2008), An exploratory study of associations between social capital and self-assessed health in Norway, *Health Econ Polic Law*, 3: 349-364.
- Johnell K, Råstam L, Lithman T et al. (2005), Low adherence with antihypertensives in actual practice: the association with social participation - a multilevel analysis, *BMC Public Health*, 5: 17.
- Kaggwa EB, Diop N, Storey JD (2008), The role of individual and community normative factors: a multilevel analysis of contraceptive use among women in union in Mali, *Int Fam Plan Perspec*, 34 (2): 79-88.
- Kahn RS, Wise PH, Kennedy BP, Kawachi I (2000), State income inequality, household income, and maternal mental and physical health: cross sectional national survey, *BMJ*, 321: 1311-1315.
- Kelsey J, Whittemore A, Evans A et al. (1996), *Methods in observational epidemiology*, Oxford, University Press.
- Kish L (1965), *Survey sampling*, New York, John Wiley & Sons.
- Kogan MD, Singh GK, Dee DL et al. (2008), Multivariate analysis of state variation in breastfeeding rates in the United States, *Am J Public Health*, 98 (10): 1872-1880.
- Krein SL, Hofer TP, Kerr EA et al. (2002), Whom should we profile? Examining diabetes care practice variation among primary care providers, provider groups and health care facilities, *Health Serv Res*, 37 (5): 1159-1180.
- Kuntsche E, Kuendig H, Gmel G (2008), Alcohol outlet density, perceived availability and adolescent alcohol use: a multilevel structural equation model, *J Epidemiol Community Health*, 62 (9): 811-816.
- Leung K, Elashoff R, Rees K et al. (1998), Hospital- and patient-related characteristics determining maternity length of stay: a hierarchical linear model approach, *Am J Public Health*, 88 (3): 377-381.
- Leyland AH, Boddy FA (1998), League tables and acute myocardial infarction, *Lancet*, 351: 555-558.
- Lippi Bruni M, Nobilio L, Ugolini C (2009), Economic incentive in general practice: the impact of pay-for-participation and pay-for-compliance programs on diabetes care, *Health Policy* (in press).
- Lochner K, Pamuk ER, Makuc D, Kennedy BP, Kawachi I (2001), State-level income inequality and individual mortality risk: a prospective multilevel study, *Am J Pub Health*, 91: 385-391.
- Longford N (1993), *Random coefficient models*, Oxford, Clarendon Press.
- Lopez R (2004), Income inequality and self-rated health in US metropolitan areas: a multi-level analysis, *Soc Sci Med*, 59: 2409-2419.
- Luo R, Giordano SH, Zhang DD et al. (2007), The role of the surgeon in whether patients with lymph node-positive colon cancer see a medical oncologist, *Cancer*, 109 (5): 975-982.
- Maas CJM, Hox, JJ (2005), Sufficient sample sizes for multilevel modeling, *Methodology*, 1 (3): 86-92.
- Manca A, Rice N, Sculpher MJ et al. (2005), Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models, *Health Econ*, 14: 471-485.
- McCullagh P, Nelder J (1989), *Generalized linear models*, London, Chapman and Hall.
- Moineddin R, Matheson F, Clazier R (2007), A simulation study of sample size for multilevel logistic regression, *BMC Medical Research Methodology*, 7: 34-44.
- Monzani E, Erlicher A, Lora A et al. (2008), Does community care work? A model to evaluate the effectiveness of mental health services, *Int J Ment Health Syst*, 2 (1): 10.
- Or Z, Wang J, Jamison D (2005), International differences in the impact of doctors on health: a multilevel analysis of OECD countries, *J Health Econ*, 24: 531-560.
- Paccagnella O (2006), Comparing vocational training courses through a discrete-time multilevel hazard model, *Stat Modelling*, 6: 119-139.
- Pelletier-Fleury N, Le Vaillant M, Hebbrecht G et al. (2007), Determinants of preventive services in general practice. A multilevel approach in cardiovascular domain vaccination in France, *Health policy*, 81: 218-227.
- Pinheiro J, Bates D (1995), Approximations to log-likelihood function in the nonlinear mixed-effects model, *Journal Computational and Graphical Statistics*, 4 (1): 12-35.
- Rice N, Jones A (1997), Multilevel models and health economics, *Health Econ*, 6: 561-575.
- Robinson WS (1950), Ecological correlations and the behaviour of individuals, *Am Sociol Rev*, 15: 321-357.

- Scott A, Shiell A (1997a), Do fee descriptors influence treatment choices in general practice? A multilevel discrete choice model, *J Health Econ*, 16: 323-342.
- Scott A, Shiell A (1997b), Analysing the effect of competition on general practitioners' behaviour using a multilevel modelling framework, *Health Econ*, 6: 577-588.
- Shenkman E, Wu S, Nackashi J et al. (2003), Managed care organizational characteristics and health care use among children with special health care needs, *Health Serv Res*, 3 (6: Part I), 1599-1624.
- Snijders T, Bosker R (1993), Standard errors and sample sizes in two-level research, *J Educ Stat*, 18 (3), 237-259.
- Snijders T, Bosker R (1994), Modelled variance in two-level models, *Sociol Method Res*, 22: 342-363.
- Snijders T, Bosker R (1999), *Multilevel analysis: an introduction to basic and advanced multilevel modeling*, Thousand Oaks (CA), Sage Publications.
- Subramanian SV, Kawachi I (2006), Whose health is affected by income inequality?, *Health & Place*, 12 (2): 141-156.
- Sundquist K, Malmström M, Johansson SE (2004), Neighbourhood deprivation and incidence of coronary heart disease: a multilevel study of 2.6 million women and men in Sweden, *J Epidemiol Community Health*, 58: 71-77.
- Verstappen W, ter Riet G, van der Weijden T et al. (2005), Variation in requests for imaging investigations by general practitioners: a multilevel analysis, *J Health Serv Res Policy*, 10 (1): 25-30.
- Von Korff M, Koepsell T, Curry S et al. (1992), Multilevel analysis in epidemiologic research on health behaviors and outcomes, *Am J Epidemiol*, 132: 1077-1082.
- Webster BS, Cifuentes M, Verma S et al. (2008), Geographic variation in opioid prescribing for acute, work-related, low back pain and associated factors: a multilevel analysis, *Am J Ind Med*, 52 (2): 162-171.
- Williams CT, Latkin CA (2007), Neighborhood socioeconomic status, personal network attributes, and use of heroin and cocaine, *Am J Prev Med*, 32 (6 Suppl): S203-210.