



UNIVERSITÀ DEGLI STUDI
DI MILANO

W P



WP | SPS 03/2013

Working paper del Dipartimento di Scienze Sociali e Politiche
Department of Social and Political Sciences Working Paper

FERRUCCIO BIOLCATI RINALDI*, SIMONE SARTI*,
ALBERTO VITALINI**

Le disuguaglianze sociali nella salute in Lombardia e in Italia (2010-2012):

un'analisi multilivello delle condizioni individuali e degli
effetti di contesto dei comuni

* Università degli Studi di Milano – Dipartimento di Scienze Sociali e Politiche

** Istituto Nazionale di Statistica – Sede territoriale per la Lombardia



Le disuguaglianze sociali nella salute in Lombardia e in Italia (2010-2012): un'analisi multilivello delle condizioni individuali e degli effetti di contesto dei comuni

Ferruccio Biolcati Rinaldi, Simone Sarti¹ e Alberto Vitalini

1.Introduzione

La salute della popolazione costituisce un tema di fondamentale importanza per le società umane. Relativamente alle scienze sociali essa riguarda diverse aree tematiche: culturale, politica ed economica, per citare le principali. Buone condizioni di salute permettono agli individui di riprodursi non solo in termini biologici, ma anche in senso culturale attraverso la formazione di identità stabili e ben adattate (Gallino 1994). Il miglioramento della salute di tutti gli esseri umani costituisce a livello mondiale uno dei più importanti obiettivi dichiarati dagli organismi internazionali, dall'ONU all'OMS. Più prosaico, ma non meno rilevante, è anche il peso economico della salute nelle moderne società. Nei paesi occidentali, per esempio, circa il dieci per cento del Prodotto Interno Lordo è destinato alla sanità, e tale quota è in tendenziale aumento².

Con il concetto di salute si possono intendere aspetti differenti dello "star bene", non solo l'assenza di malattie o disabilità, ma anche il benessere psico-fisico in senso più generale in termini di qualità della vita. La salute può quindi essere declinata operativamente attraverso molteplici indicatori: i tassi di mortalità, la presenza di patologie, la salute auto-percepita, i livelli di stress, l'esposizione a certi fattori di rischio, solo per richiamarne alcuni³. Tali indicatori, le cui osservazioni sono tradotte in "stati di salute", presentano su una popolazione un'elevata eterogeneità: banalmente alcuni soggetti mostreranno una salute migliore, altri una salute peggiore.

L'individuazione delle condizioni sociali associate a tale eterogeneità rappresenta l'obiettivo più generale della ricerca empirica che tratta il tema della salute dal punto di vista sociologico.

Quando la salute si mostra associata a condizioni sociali che per una qualsiasi ragione sono riconducibili a forme di disuguaglianza allora si parla di disuguaglianze sociali nella salute. In tal senso la variabilità negli stati di salute è vista come il riflesso di svantaggi strutturali insiti nell'organizzazione sociale, dalla stratificazione del sistema educativo fino all'assetto del mercato del lavoro (Mackenbach *et al.* 2003; Bartley 2003; Cavelaars *et al.* 1998; Link e Phelan 1995; Fox 1989; Whitehead 1987). Più rigorosamente possiamo definire le disuguaglianze sociali nella salute come "disparità oggettive e sistematiche riguardanti il possesso di risorse sociali, economiche e culturali e la connessa capacità di utilizzare queste risorse al fine di massimizzare la propensione a una piena efficienza psico-fisica del corpo. Queste disparità derivano dall'azione di individui o di gruppi o dall'operare di dispositivi selettivi, comportamentali, strutturali e ambientali, orientati

¹ Osservatorio e Metodi per la Salute – Università di Milano-Bicocca.

² Fonte Banca Mondiale: <http://data.worldbank.org/>

³ Per un inquadramento di questi fenomeni si veda Istat: Sabbadini *et al.* (2007): http://www3.istat.it/salastampa/comunicati/non_calendario/20070302_00/

all'ottenimento di una migliore qualità della vita e più in generale alla sopravvivenza biologica“ (Sarti 2006: 320).

2. Interrogativo di ricerca

Quando si indagano le disuguaglianze sociali nella salute occorre tenere ben presente che nella relazione tra condizioni sociali e salute possono presentarsi diversi confondenti, ovvero variabili che, una volta prese in considerazione, possono annullare o comunque ridurre il ruolo delle condizioni sociali.

Questi “confondenti” possono essere classificati in due ambiti tra loro interagenti: biologico e contestuale⁴.

Il primo sottosta alle caratteristiche genetiche degli individui ossia il loro patrimonio genetico può intrinsecamente essere portatore di maggiori propensioni a certe malattie. Non meno importanti nel determinare la salute sono le caratteristiche ecologiche di un'area, per cui esiste un effetto presente per tutti i soggetti che risiedono o appartengono ad un certo contesto, indipendentemente dalle loro caratteristiche individuali, ed assente (o differente) per i soggetti che non vi appartengono. Dunque identificare un'associazione tra alcune variabili che costituiscono *proxy* delle risorse sociali (come tipicamente sono il titolo di studio o la condizione occupazionale o la classe sociale) non è sufficiente a descrivere le disuguaglianze sociali nella salute: occorre tenere sotto controllo alcune variabili biologiche chiave. In assenza di informazioni sul patrimonio genetico generalmente si utilizzano il sesso e l'età anagrafica, che costituisce la determinante più rilevante della cattiva salute.

D'altra parte, quando disponibile, occorre tenere anche in considerazione l'area territoriale cui gli individui appartengono.

L'oggetto di ricerca più innovativo che presentiamo in questo studio è il livello ecologico che ci è permesso tenere sotto controllo grazie ai dati resi disponibili dall'ISTAT: il livello comunale. Attraverso adeguati modelli statistici multivariati (multilivello) indagheremo l'associazione tra le condizioni sociali degli individui sulla salute al netto di alcuni confondenti tra cui l'area di residenza (Goldstein 1995).

Questo costituisce un tema assai rilevante nella letteratura sociologica ed epidemiologica. Le evidenze empiriche mostrano infatti risultanti discordanti sul ruolo dei contesti territoriali. Mentre è assodato il ruolo delle condizioni sociali secondo cui a titoli di studio più elevati, a condizioni occupazionali migliori, all'appartenenza a classi sociali più vantaggiose, corrisponde sistematicamente una migliore salute, in termini di una più alta speranza di vita, di una minore propensione alle patologie, e di una migliore salute percepita; rimane più discusso il ruolo dei contesti territoriali. Il ruolo da essi esercitato appare decisamente più articolato: se diversi studi epidemiologici mostrano che la presenza di inquinamento o di condizioni climatiche avverse (come ad esempio il particolato atmosferico o le ondate di calore) sono correlate a peggiori stati di salute (Bell *et al.* 2013), più contraddittori sono gli studi sociologici: ad esempio alcuni di essi mostrano che una maggiore disuguaglianza produrrebbe in generale peggiori condizioni di salute, altri mettono in discussione questi risultati, anche in funzione delle diverse nazioni prese a riferimento (Wilkinson e Pickett 2009; Wilkinson 2002; Tremblay *et al.* 2002).

⁴ A questi potremmo aggiungere i comportamenti individualizzati o meramente idiosincratici (auto-determinati sulla base di libere scelte) e l'effetto del caso, ma entrambi, entro i modelli statistici impiegati, rimangono nella parte di varianza non spiegata (vedi Sarti *et al.* 2011).

La definizione stessa dei contesti territoriali pare modificare gli esiti degli studi, tendenzialmente minore è la dimensione dei contesti impiegati (idealmente fino a scendere a livello di quartieri o isolati) maggiore è la forza probante del modello preso in esame. In quanto la micro-mappatura delle condizioni contestuali permette di escludere effetti aggregati che all'aumentare delle dimensioni del contesto possono regredire sulla media e quindi risultare non significativi.

In Italia la letteratura sull'argomento è limitata. Citiamo qui gli studi più recenti di Marinacci *et al.* (2010), Lucchini *et al.* (2009) e il lavoro di Piombo (2013). Tutti questi studi sono concordi nel ricondurre l'eterogeneità nella salute ad effetti legati alle condizioni individuali. In altre parole, una volta controllata la condizione sociale per sesso, età, titolo di studio e condizione occupazionale, l'eterogeneità nella salute legata ai contesti risulta estremamente modesta. Ciò significa che la salute pare determinata da caratteristiche individuali risultato di effetti di composizione e non da caratteristiche di area. In altre parole, se in certe aree territoriali la popolazione sta peggio in termini di salute ciò è mediamente dovuto al fatto che le condizioni sociali individuali di quella popolazione sono mediamente peggiori. Tuttavia gli studi succitati impiegano come unità contestuale la regione, oppure un numero limitato di comuni. In tal senso un livello ecologico più dettagliato o più esteso sul piano geografico potrebbe condurre a risultati assai differenti, come ad esempio rilevato in alcuni studi (Schaefer *et al.* 2010; Diez Roux 2001; Marinacci *et al.* 2004).

In questo lavoro siamo dunque interessati a studiare la relazione tra salute e condizioni sociali (rilevate attraverso proxy delle risorse socio-economiche di cui gli individui dispongono) tenendo sotto controllo i principali confondenti demografici ed ecologici (a livello di comune di residenza).

Il nostro interrogativo di ricerca è quindi il seguente: l'eterogeneità geografica dipende o meno da effetti di composizione? Ossia le disuguaglianze di salute sono maggiormente influenzate dalle caratteristiche individuali o dalla diversa appartenenza geografica?

3.Dati e metodo

I dati utilizzati provengono dalle indagini Istat "Multiscopo sulle famiglie: aspetti della vita quotidiana", riferite agli anni 2010, 2011 e 2012. Le tre edizioni dell'indagine sono state cumulate al fine di raggiungere una numerosità campionaria adeguata per le analisi (Biolcati-Rinaldi e Vezzoni 2012: 49-51). Ogni anno sono state intervistate circa 20.000 famiglie campione e 50.000 individui, selezionati secondo un disegno di campionamento a più stadi, con stratificazione delle unità di primo stadio rappresentate dai comuni. All'interno di ciascun comune estratto sono selezionate, in maniera sistematica, dalle liste anagrafiche alcune famiglie. A tutti i componenti delle famiglie estratte è somministrata l'intervista.

Informazioni a livello comunale sono state ricavate dai dati del Censimento della popolazione 2011 (disponibili nel data warehouse I.Stat) ed estratte dal sistema informativo Istat "Atlante dei comuni italiani".

Sono stati selezionati individui in età compresa tra i 27 e i 69 anni. Questa scelta è legata a ragioni di tipo teorico: indagare individui che hanno completato il corso degli studi ed evitare, al contempo, l'introduzione nell'analisi di distorsioni dovute all'effetto della mortalità (Willson *et al.* 2007).

I dati sono quindi organizzati su tre livelli gerarchici: individui, famiglie e comuni. Le numerosità per la Lombardia sono rispettivamente: 7.014, 4.054 e 196. Per l'Italia sono: 83.375, 47.102 e 1.758⁵.

3.1 Variabile dipendente

La salute percepita è la variabile dipendente utilizzata nelle analisi. Nell'indagine "Multiscopo sulle famiglie - aspetti della vita quotidiana" la salute percepita è rilevata ponendo all'intervistato la seguente domanda: "Come va in generale la sua salute?". L'intervistato può scegliere, tra cinque modalità di risposta, quella più adatta alla sua condizione: "Molto bene", "Abbastanza bene", "Né bene, Né male", "Male" e "Molto male".

Pur nella sua semplicità la condizione di salute percepita e dichiarata dall'individuo, rilevata con formulazioni simili a quelle utilizzate nell'indagine multiscopo, è risultata, sia in ambito internazionale che nazionale, non solo un buon indicatore della reale condizione di salute fisica e mentale, ma è anche ritenuto un predittore della mortalità negli anni successivi (Jylha 2009; Singh-Manoux *et al.* 2007; Egidi e Spizzichino 2006; Fayers e Sprangers 2002; Idler e Benyamini 1997).

3.2 Variabili indipendenti

Livello individuale e familiare

Sono state testate a livello individuale variabili di tipo demografico e sociale: genere, età, educazione, stato occupazionale. L'età e l'educazione sono state inserite nei modelli come variabili sia continue sia categoriali nell'ipotesi che il loro impatto sulla salute percepita non fosse lineare attraverso il corso di vita⁶. Tuttavia nei modelli finali sono state inserite come variabili metriche centrate sulla media dal momento che gli effetti riscontrati sulla salute in entrambi i casi approssimano un andamento lineare.

È stato anche valutato l'impatto sulla salute della classe sociale familiare⁷. Quest'ultima è stata costruita associando, in un primo momento, ad ogni membro una classe sociale individuale sulla base dell'ultima occupazione dichiarata secondo i tradizionali schemi di classe impiegati in sociologia (Schizzerotto 2002; Erikson e Goldthorpe 1992): imprenditori, dirigenti, direttivi, quadri e liberi professionisti sono stati codificati come "Borghesia", impiegati e intermedi come "Classe Media Impiegatizia", lavoratori in proprio e coadiuvanti come "Piccola borghesia/Lavoratori autonomi", capi operai, operai, apprendisti, soci cooperativi, lavoratori a domicilio, a contratto co.co.co, a progetto e occasionali come "classe operaia". Dopodiché è stato applicato un criterio di dominanza attribuendo ad ogni famiglia la classe più elevata presente tra i suoi componenti con la seguente gerarchia: borghesia, classe media impiegatizia, piccola borghesia (lavoratori autonomi), classe operaia. I soggetti che non ricadevano in nessuna delle categorie precedenti sono stati inseriti nella categoria residuale "altro" (Erikson 1984; Pisati 2000).

⁵ Nel campione a livello italiano all'interno delle famiglie ci sono in media 1,8 individui all'interno dei comuni sono in media 47,4. In Lombardia i rispettivi valori medi sono 1,7 e 35,8.

⁶ Il titolo di studio è stato trasformato in una variabile metrica: ad ogni livello formale di istruzione è stato associato il numero di anni di scuola necessari al suo conseguimento.

⁷ Sono state inoltre considerate il numero di componenti della famiglia e alcune differenti tipologie familiari. L'inserimento di queste ulteriori variabili di controllo non muta le stime degli altri regressori considerati. I modelli completi sono disponibili su richiesta agli autori.

L'utilizzo della classe sociale familiare congiuntamente alle caratteristiche individuali (disposte su due differenti livelli) va considerato un elemento di originalità del presente lavoro che si basa sulla convinzione che le diseguaglianze di salute dipendano non solo da risorse strettamente personali ma anche familiari, ossia anche da risorse di vario tipo (economiche, abitative, cognitive, legate agli stili di vita) che l'individuo può non possedere direttamente ma condividere con gli altri membri. Si vedano in Tabella 1 le distribuzioni delle variabili di livello individuale e familiare impiegate nei successivi modelli multivariati.⁸

Livello comunale

Sono state prese in considerazione diverse variabili a livello comunale al fine di spiegare la variabilità territoriale residua della salute percepita, in caso di sua significativa presenza: latitudine, tipologia di comune ("montagna interna", "montagna litoranea", "collina interna", "collina litoranea", "pianura"), densità abitativa, densità abitativa per abitanti equivalenti⁹, reddito medio imponibile delle persone fisiche ai fini delle addizionali all'Irpef (solo per la regione Lombardia).

⁸ Si vedano in Appendice A1 le distribuzioni delle variabili originali titolo di studio, condizione e posizione occupazionale.

⁹ L'abitante equivalente è convenzionalmente definito come la quantità di carico inquinante biodegradabile prodotto ed immesso in fognatura da un abitante stabilmente residente nel centro urbano nell'arco della giornata.

Tab.1. Distribuzioni percentuali delle variabili genere, età all'intervista, titolo di studio, condizione occupazionale, classe sociale familiare e salute percepita e valori medi dell'età e degli anni di scolarità.¹⁰

	Lombardia	Italia
Genere		
Maschi	49,0	48,9
Femmine	51,0	51,1
Età all'intervista		
27-39 anni	29,0	28,6
40-49 anni	27,8	27,1
50-59 anni	22,7	23,6
60-69 anni	20,5	20,7
Età media all'intervista	47,6	47,7
Titolo di studio		
Licenza elementare	12,0	14,1
Licenza media	42,8	40,1
Diploma di scuola sup.	30,8	31,8
Titolo universitario	14,4	14,1
Anni di scolarità medi	10,8	10,6
Condizione occupazionale		
Occupato	63,6	56,6
In cerca di occupazione	5,6	10,0
Inattivo	30,8	33,4
Classe sociale familiare		
Bor - Borghesia	24,7	21,6
CMI - Classe Media Impiegatizia	38,1	34,6
AUT - Lavoratori autonomi	9,9	13,4
OPE - Classe operaia	26,4	27,3
Altro	0,9	3,1
Salute Percepita		
1 Peggior	0,3	0,6
2	2,4	3,5
3	25,2	24,8
4	58,6	56,5
5 Migliore	13,5	14,7
N	7.014	83.375

3.3 Tecniche di analisi

Per le analisi sono stati utilizzati modelli di regressione multilivello nell'ipotesi che la salute sia il risultato di fattori che interagiscono a livello sia d'individuo sia di contesto. In letteratura lo studio dei modelli di regressione multilivello è dovuto alla necessità di individuare strumenti statistici appropriati per analizzare dati con strutture gerarchiche: cioè strutture in cui le unità individuali sono contenute in aggregati di livello superiore.

¹⁰ Nelle elaborazioni successive sono stati impiegati i casi non pesati, poiché è risultato che l'uso o meno dei coefficienti di riporto all'universo non influenza la validità dei risultati sostanziali ed introduce una complessità non necessaria nelle elaborazioni. Alla stessa conclusione è giunta Piombo (2013) in un lavoro, simile al nostro, nel quale sono stati utilizzati i dati dell'indagine campionaria sulle "Condizioni di Salute e Ricorso ai Servizi Sanitari 2004-2005" (<http://www.istat.it/it/archivio/5471>).

I modelli di regressione multilivello si adattano particolarmente bene alla struttura gerarchica dei dati campionari a disposizione: gli intervistati sono, infatti, distribuiti su tre livelli: individuo, famiglia, comune. Il principale vantaggio dei modelli di regressione multilivello, nel quadro di questo contributo, è quello di distinguere tra effetti delle caratteristiche individuali, familiari e comunali e di testare eventuali effetti di interazione tra i livelli. Per una più completa descrizione dei vantaggi statistici delle regressioni multilivello rispetto a quelle tradizionali si rimanda a Goldstein (1995).

In questo contributo sono presentate le stime di regressioni multilivello lineari, calcolate utilizzando il programma MLwiN.¹¹ La variabile dipendente - salute percepita - è considerata come una proxy di una variabile metrica; per questo motivo si è attribuito il valore 1 alla modalità di risposta “Molto male”, 2 a “Male”, 3 a “Né bene, Né male”, 4 a “Abbastanza bene” e, infine, 5 a “Molto bene”¹².

La strategia d’implementazione dei modelli presentati è di tipo *bottom-up* e prevede che diversi modelli siano sviluppati in modo incrementale (Hox 2010).

Il primo modello, solitamente chiamato “nullo” o “vuoto” o “con solo l’intercetta” non contiene variabili indipendenti e permette di scomporre la variabilità delle risposte tra i diversi livelli di analisi. A questo modello base sono in seguito aggiunte variabili indipendenti ai diversi livelli (individuale, familiare, comunale) per valutare il loro specifico contributo alla previsione del livello di salute percepita.

¹¹ Le analisi sono state ripetute con STATA 12.

¹² Diversi tipi di trasformazioni (ad esempio logaritmica, radice quadrata) sono stati applicati a questa variabile per tentare di normalizzare la sua distribuzione che è asimmetrica. Nessuna trasformazione ha mostrato risultati apprezzabili, pertanto si è deciso di utilizzare la variabile non trasformata.

Inoltre nella consapevolezza dei problemi statistici legati all’utilizzo di una variabile ordinale come proxy di una metrica, sono state calcolate regressioni multilivello multinomiali con le stesse variabili indipendenti, dove la variabile dipendente -salute percepita- è stata introdotta come ordinale. I risultati dei due tipi di regressione, in termini di parametri, sono congruenti. Questo risultato ha confermato empiricamente la decisione di utilizzare la variabile ordinale -salute percepita- come una proxy di una variabile metrica. L’utilizzo delle regressioni multilivello lineari, pur forzando la natura ordinale della variabile dipendente, è stato preferito rispetto a quelle multinomiali perché le prime a differenza delle seconde permettono un semplice calcolo della devianza dei modelli (valore che riassume l’adattamento del modello ai dati) con la conseguente possibilità di fare inferenze sulla specificazione dei modelli generati.

4. Risultati

I principali risultati della analisi condotte sono riportati nelle Tabelle 2 e 3.¹³ Le analisi sono state prima svolte per l'Italia nel suo complesso e successivamente replicate per la sola Lombardia, dato il focus locale dal quale il progetto di ricerca ha inizialmente mosso i passi.

Il dibattito al quale il paper vuole contribuire riguarda la natura individuale piuttosto che contestuale (qui intesa come: comunale) delle disuguaglianze di salute: a tal fine è stato costruito un dataset multilivello che permette di distinguere gli effetti individuali da quelli contestuali. Va evidenziato come sia stata qui utilizzata una *household survey* (Indagine multiscopo sulle famiglie – Aspetti della vita quotidiana) che consente di isolare anche il livello familiare.

La risposta al nostro principale obiettivo di ricerca viene dal modello 1.0 della Tabella 2. Tale modello contiene la sola intercetta e permette di stimare la proporzione di varianza attribuibile ai vari livelli. Più di due terzi della varianza della variabile relativa alla salute percepita è attribuibile al livello individuale (70,1%) e circa un quarto al livello familiare (25,6%). Risulta quindi minima la proporzione di varianza attribuibile al livello contestuale (comunale): 4,3%. Questi risultati confermano la rilevanza del livello individuale e familiare a scapito di quello comunale: le disuguaglianze di salute sembrano quindi riconducibili a effetti di composizione piuttosto che a genuini effetti contestuali (legati alla qualità ambientale e demografica dei diversi comuni). A questo punto risulta di scarsa utilità ricercare caratteristiche contestuali a livello comunale capaci di spiegare le eterogeneità territoriali. Abbiamo quindi tralasciato di inserire nei modelli le variabili ecologiche selezionate in partenza, come altimetria, latitudine, urbanizzazione, eccetera (si veda il paragrafo 3.2).

Nei modelli successivi vengono inserite una serie di variabili indipendenti per cercare di rendere conto delle varianze ai vari livelli. Facendo riferimento al modello 1.3 della Tabella 1, vediamo che sono state inserite tre variabili a livello individuale che hanno un effetto significativo sull'indicatore. I maschi hanno una salute percepita maggiore di quasi un decimo di punto (0,09) rispetto alle femmine: si ricorda che la scala dell'indicatore va da 1 a 5, ma, essendo la sua distribuzione fortemente concentrata attorno alla media, si tratta di un effetto certamente apprezzabile. L'età e la scolarità hanno effetti di entità simile ma di segno opposto: la salute percepita aumenta al diminuire dell'età e all'aumentare degli anni di scolarità. L'effetto è pari a -0,02 per l'età e 0,02 per la scolarità: considerando dieci anni e non uno solo, l'effetto è pari a circa due decimi di punto. A livello familiare è stata inserita la classe sociale i cui effetti significativi sostanzialmente confermano l'usuale gerarchia. Rispetto alla categoria della classe operaia, gli individui appartenenti a famiglie di estrazione borghese hanno un vantaggio competitivo di più di un decimo di punto (0,11), mentre è simile (poco maggiore di mezzo punto) l'effetto per la classe media impiegatizia (0,07) e per la piccola borghesia (0,06). Infine, la categoria "altro" ha uno svantaggio di circa mezzo punto (-0,05) sempre rispetto alla categoria della classe operaia.

Nel modello 1.4 sempre della Tabella 1, è stata aggiunta la condizione occupazione a livello individuale. Gli individui in cerca di occupazione e inattivi hanno una salute percepita inferiore di circa un decimo di punto rispetto agli occupati (i coefficienti di regressione sono rispettivamente -0,09 e -0,11). L'inserimento di questa variabile comporta una contrazione degli effetti relativi al genere e alla classe operaia.

Nella Tabella 3 sono riportati i risultati delle analisi focalizzate sulla Lombardia. Come prevedibile, c'è qualche variazione nella scomposizione della varianza (modello 1.0). Diminuisce infatti la proporzione di varianza a livello comunale (2,1%) a favore del livello individuale (73,1%). Diminuisce anche la varianza a livello familiare (24,8%), ma decisamente meno in termini relativi.

Le conclusioni tratte a livello italiano relativamente al dibattito sulla natura individuale piuttosto che contestuale delle disuguaglianze di salute sono quindi confermate a livello regionale (Lombardia). I valori dei coefficienti di regressione per il sottocampione lombardo sono sostanzialmente analoghi a quelli del campione nazionale, per cui si ritiene superfluo qui commentarli.

¹³ Le equazioni di stima dei parametri dei modelli sono riportate nelle figure A2 e A3 in appendice.

5. Conclusioni

In questo studio ci siamo chiesti se le disuguaglianze di salute sono maggiormente influenzate dalle caratteristiche individuali o dalla diversa appartenenza geografica. I risultati delle nostre analisi mostrano, come rilevato in altre ricerche, che l'eterogeneità nei livelli di salute appare imputabile più ad effetti di composizione che a effetti attribuibili al contesto territoriale (Piombo 2013; Marinacci et al. 2010; Lucchini et al. 2009). L'esistenza di una forte variabilità della salute percepita è quindi parzialmente prevedibile sulla base dell'età, del genere, del livello di istruzione, della classe sociale familiare e della condizione occupazionale. Le variazioni territoriali seppur statisticamente significative sono di scarsa rilevanza sostanziale. Sintetizzando i risultati in una frase potremmo affermare che: "È ciò che sei e ciò che fai, non dove vivi, che ha una maggiore influenza sullo stato della tua salute percepita" (Tremblay *et al.* 2002).

I risultati di questo lavoro vanno comunque letti alla luce di tre importanti puntualizzazioni.

In primo luogo occorre ricordare che potrebbero esistere effetti di area nel momento in cui si considerano contesti territorialmente più piccoli. Il contesto sociale che abbiamo considerato è quello comunale, che dovrebbe presentare caratteristiche socio-economiche omogenee al suo interno. In ogni caso nei comuni di elevata ampiezza è possibile che l'eterogeneità dei diversi aspetti socio-economici sia elevata. Uno studio sui fattori che influenzano il livello di salute che consideri come contesto le sezioni di censimento (aree per le quali saranno disponibili i dati nel prossimo anno) potrebbe portare a conclusioni diverse da quelle esposte in questo testo.

In secondo luogo, l'adozione di modelli multilivello è una pratica consolidata nella letteratura per valutare separatamente gli effetti delle caratteristiche individuali, familiari e di contesto sul livello di salute (Schaefer-McDaniel *et al.* 2010). Tuttavia essa non permette di chiarire definitivamente la relazione fra effetti di tipo contestuale e individuale. Solitamente nei modelli di regressione multilivello, si aggiungono al modello "nullo" le variabili di livello individuale e/o familiare. Se la variabilità a livello di contesto risulta a questo punto statisticamente non significativa si può affermare che le differenze in termini di salute sono dovute ad effetti di composizione e non ad effetti di contesto. La conclusione, pur non essendo in contraddizione con i dati, può non essere corretta. Le variabili individuali introdotte potrebbero essere delle variabili intervenienti: in altre parole gli attributi individuali o familiari potrebbero essere essi stessi plasmati da caratteristiche dell'area in cui si vive; ad esempio, la salute è influenzata dalla classe sociale ma quest'ultima può essere condizionata dalle caratteristiche del mercato del lavoro locale.¹⁴

Infine, in riferimento al tema della salute, diventa particolarmente rilevante la questione dell'eterogeneità non osservata legata al patrimonio genetico degli individui: tale patrimonio ha certamente una variabilità individuale la cui gran parte è però riconducibile all'ereditarietà familiare. Non è l'obiettivo di questo studio stimare tale componente genetica a livello individuale e/o familiare, ma si tratta di un aspetto da considerare (Lucchini *et al.* 2011).

¹⁴ Per una disamina critica di questo aspetto, e più in generale, delle problematiche legate allo studio degli effetti di composizione e contestuali con l'utilizzo di modelli di regressione si veda l'articolo di Macintyre *et al.* (2002).

Tabella 2. Italia: modelli multivello di regressione lineare sulla salute percepita (1-5): stime di regressione e deviazioni standard (N=83.375).¹⁵

	<i>Modello 1.0</i>	% di varianza	<i>Modello 1.1</i>	% di varianza	<i>Modello 1.2</i>	% di varianza	<i>Modello 1.3</i>	% di varianza	<i>Modello 1.4</i>	% di varianza
Intercetta ⁺	3,80 (0,01)		3,75 (0,01)		3,76 (0,01)		3,71 (0,01)		3,77 (0,01)	
PRIMO LIVELLO										
Genere (M)			0,09 (0,00)		0,09 (0,00)		0,09 (0,00)		0,07 (0,00)	
Età (27-69 anni) centrata			-0,02 (0,00)		-0,02 (0,00)		-0,02 (0,00)		-0,02 (0,00)	
Anni di scolarità (0-21) centrati					0,02 (0,00)		0,02 (0,00)		0,02 (0,00)	
<i>Condizione occupazionale:</i>										
Occupato									0 ^a	
In cerca di occupazione									-0,09 (0,01)	
Inattivo									-0,11 (0,01)	
SECONDO LIVELLO										
<i>Classe sociale</i>										
Borghesia							0,12 (0,01)		0,11 (0,01)	
CMI							0,07 (0,01)		0,06 (0,01)	
PB							0,06 (0,01)		0,05 (0,01)	
Altro							-0,05 (0,02)		-0,01 (0,02)	
OPE							0 ^a		0 ^a	
Random:										
Varianza individui	0,389	70,1%	0,337	68,9%	0,336	70,1%	0,335	70,2%	0,334	70,5%
Varianza famiglie	0,142	25,6%	0,128	26,2%	0,120	25,1%	0,119	24,9%	0,118	24,9%
Varianza comuni	0,024	4,3%	0,024	4,9%	0,023	4,8%	0,023	4,8%	0,022	4,6%
<i>IGLS Deviance (-2Ln(L))</i>	182.365		171.185		169.817		169.566		169.190	

⁺ *Random parameter*

¹⁵ Note: Per testare la significatività statistica di un parametro fisso si può utilizzare il test parametrico z con distribuzione normale. Il valore della statistica z per un dato coefficiente è dato dal rapporto fra il valore del coefficiente e quello dell'errore standard (scritto tra parentesi in tabella). Le celle ombreggiate contengono coefficienti statisticamente non significativi al livello dello 0,05. Per confrontare i modelli tra loro si calcola invece il test LR (Likelihood ratio). Il valore della statistica LR è dato dalla differenza, in valore assoluto, della IGLS Deviance dei due modelli da testare. La statistica test LR ha una distribuzione chi-quadro con gradi di libertà pari al numero di parametri stimati in più nel modello più complesso.

Tabella 3. Lombardia: modelli multivello di regressione lineare sulla salute percepita (1-5): stime di regressione e deviazioni standard (N=7.014).¹⁶

	<i>Modello 1.0</i>	% di varianza	<i>Modello 1.1</i>	% di varianza	<i>Modello 1.2</i>	% di varianza	<i>Modello 1.3</i>	% di varianza	<i>Modello 1.4</i>	% di varianza
Intercetta ⁺	3,82 (0,01)		3,77 (0,01)		3,78 (0,01)		3,72 (0,02)		3,77 (0,02)	
PRIMO LIVELLO										
Genere (M)			0,11 (0,01)		0,11 (0,01)		0,12 (0,01)		0,09 (0,02)	
Età (27-69 anni) centrata			-0,02 (0,00)		-0,02 (0,00)		-0,02 (0,00)		-0,02 (0,00)	
Anni di scolarità (0-21) centrati					0,02 (0,00)		0,02 (0,00)		0,02 (0,00)	
<i>Condizione occupazionale:</i>										
Occupato									0 ^a	
In cerca di occupazione									-0,03 (0,03)	
Inattivo									-0,10 (0,02)	
SECONDO LIVELLO										
<i>Classe sociale</i>										
Borghesia							0,14 (0,03)		0,13 (0,03)	
CMI							0,04 (0,02)		0,04 (0,02)	
PB							0,06 (0,03)		0,05 (0,03)	
Altro							-0,06 (0,09)		-0,04 (0,09)	
OPE							0 ^a		0 ^a	
Random:										
Varianza individui	0,354	73,1%	0,317	72,7%	0,317	73,7%	0,317	74,2%	0,315	73,9%
Varianza famiglie	0,120	24,8%	0,109	25,0%	0,103	24,0%	0,101	23,7%	0,101	23,7%
Varianza comuni	0,010	2,1%	0,010	2,3%	0,010	2,3%	0,009	2,1%	0,010	2,3%
<i>IGLS Deviance (-2Ln(L))</i>	14.540		13.793		13.709		13.681		13.660	

⁺Random parameter

¹⁶ Per la lettura della significatività delle stime si veda la nota precedente.

RIFERIMENTI:

- Bartley, M. (2003) Commentary: relating social structure and health, in *International Journal of Epidemiology*, n.32, pp.958-960
- Bell,M.L., Zanobetti,A. e Dominici,F. (2013) Evidence on Vulnerability and Susceptibility to Health Risks Associated With Short-Term Exposure to Particulate Matter: A Systematic Review and Meta-Analysis, in *American Journal of Epidemiology*. (2013) 178 (6): 865-876 first published online July 25, 2013 doi:10.1093/aje/kwt090
- Biolcati-Rinaldi, F. e Vezzoni, C. (2012) *L'analisi secondaria nella ricerca sociale*, Il Mulino, Bologna
- Cavelaars,A.E.J.M., Kunst,A.E., Geurts,J.J.M., Helmer,U., Lundberg,O., Mielck,A., Matheson,J., Mizrahi,Ar., Mizrahi,A., Rasmussen,N., Spuhler,T., e Mackenbach, J.P. (1998) Morbidity differences by occupational class among men in seven European countries: an application of the Erikson-Goldthorpe social class scheme, in *International Journal of Epidemiology*, n.27, pp.222-230
- Diez Rouz A.V. (2001) Investigating neighborhood and area effects on health, in *American Journal of Public Health*, 91,11, pp.1783-1789
- Egidi V., Spizzichino D. (2006), "Perceived health and mortality: a multidimensional analysis of ECHP Italian data", *Genus*, LXII, n. 3-4
- Erikson, R. (1984) Social Class of men, women and families, in *Sociology*, XVIII, pp.500-514
- Erikson, R. e Goldthorpe, JH. (1992) *The Constant flux*, Oxford: Clarendon Press, 1992
- Fayers,P.M. e Sprangers, M.A.G. (2002), "Understanding self-rated health", in *The Lancet*; Jan 19, 2002; 359, 9302
- Fox, A.J. (1989) *Health inequalities in European countries*, Aldershot, Gower
- Gallino, L., (1994) *Manuale di Sociologia*, UTET, Torino
- Gargiulo, L. e Sebastiani, G. (2001) *Le condizioni di salute della popolazione. Indagine Multiscopo sulle famiglie. Condizioni di salute e ricorso ai servizi sanitari. Anni 1999-2000*, Roma, Istat.
- Goldstein, H. (1995) *Multilevel statistical models*, Halsted Press, New York
- Hox, J.J. (2010) *Multilevel analysis: Techniques and applications*, Taylor & Francis
- Idler E.L., Benyamini Y. (1997), "Self-rated health and mortality: a review of twenty-seven community studies", in *Journal of health and social behaviour*, 1997, 38:21-37
- JylhaYLHA, M. (2009) What is self-rated health and why does it predict mortality? towards a unified conceptual model , in in *"Social Science and Medicine"*, 69, pp. 307-316
- Link, B.G. e Phelan, J. (1995) Social Conditions As Fundamental Causes of Disease, in *Journal of Health and Social Behavior*, Vol.35, Extra Issue: Forty Years of Medical Sociology: The State of the Art and Directions for the Future, pp. 80-94
- Lucchini, M., Della Bella, S. and Pisati, M. (2011) The weight of the genetic and environmental dimensions in the inter-generational transmission of educational success, in *European Sociological Review*, doi: 10.1093/esr/jcr067
- Lucchini, M., Sarti S. e Tognetti, M. (2009) I welfare regionali e le differenze territoriali nelle disuguaglianze di salute, in *Dimensioni della disuguaglianza in Italia: povertà, salute, abitazione*, a cura di A.Brandolini, C.Saraceno e A.Schizzerotto, Bologna, Il Mulino, pp.165-189
- Macintyre S, Ellaway A, Cummins S. (2002) Place effects on health: how can we conceptualise, operationalise and

measure them?, in *Social Science & Medicine*, 2002;55:125-139

Mackenbach, J., Bos, V., Andersen, O., Cardano, M., Costa, G., Harding, S., Reid, A., Hemstrom, O., Valkonen, T., e Kunst, A. (2003) Widening socioeconomic inequalities in mortality in six western European countries, in *International Journal of Epidemiology*, n.32, pp.830-837

Marinacci, C., Ferracin, E., Landriscina, T., Cislighi, C., Gargiulo, L. e Costa, G. (2010) Differenze geografiche o differenze sociali, in *Rapporto OsservaSalute 2010*, pp.473-484

Marinacci, C., Spadea, T., Biggeri, A. Demaria, M., Caiazzo, A. e Costa, G. (2004) The role of individual and contextual socioeconomic circumstances on mortality: analysis of time variations in a city of North-West Italy, in *Journal of Epidemiology Community*, 58, pp. 199-207

Piombo, S. (2013) *Multilevel analysis in household surveys: an application to health condition data*, Tesi di dottorato

Pisati, M. (2000) *La Mobilità Sociale*, il Mulino, Bologna

Sabbadini, L.L., Gargiulo, L. e Sebastiani, G. (2007) *Le condizioni di salute della popolazione, Condizioni di salute, fattori di rischio e ricorso ai servizi sanitari*, Anno 2005

Sarti S. (2006) La classe sociale, in buona salute, in *POLIS*, XX, 3, Bologna, Il Mulino, pp.317-346

Sarti, S., Della Bella, S., Lucchini M., e Tognetti, M. (2011) Le disuguaglianze sociali nella salute: una riflessione sulle basi dati e sugli indicatori attualmente impiegati in letteratura , in *Rassegna Italiana di Sociologia*, 2011, 4, pp.681-702

Schaefer-McDaniel, N. O'Brien Caughy, M., O'Campo, P., Gearey, W. (2010) Examining methodological details of neighbourhood observations and the relationship to health: A literature review, in *Social Science & Medicine*, Volume 70, Issue 2, Pages 277-292, ISSN 0277-9536

Schizzerotto, A. (2002) *Vite ineguali. Disuguaglianze e corsi di vita nell'Italia contemporanea*, (a cura di), Il Mulino, Bologna

Singh-Manoux A., Guéguen A., Martikainen P., Ferrie J., Marmot M., Shipley M. (2007) "Self-Rated Health and Mortality: Short and Long Term Associations in the Whitehall II Study", *Psychosomatic Medicine*, 69:138-143

Tremblay, S., Ross, N.A. e Berthelot, J.M. (2002) Regional socio-economic context and health, in *Health Reports*, 13 (2002), pp. 1-12

Whitehead, M. (1987) *The Health Divide: Inequalities in Health in the 1980's*, Health Education Authority, London

Wilkinson, R. (2002) Commentary: liberty, fraternity, equality, in *International Journal of Epidemiology*, 31: pp.538-543

Wilkinson, R.G. e Pickett, K.E. (2009) Income inequality and social dysfunction, in *Annual Review of Sociology*, 35:493-511

Willson, A.E., Shuey, K.M. e Elder Jr., G.H. (2007) *Cumulative Advantage Processes as Mechanisms of Inequality in Life Course Health*, in «*American Journal of Sociology*», vol. 112, n. 6, pp. 1886-1924

APPENDICE:

Tabella A1. Distribuzioni di frequenza delle variabili originali: titolo di studio, condizione e posizione occupazionale.

Titolo di studio	Italia			Lombardia		
	Freq.	%	% valida	Freq.	%	% valida
1 Dottorato di ricerca	670	.8	.8	55	.8	.8
2 Laurea di 4 anni o più,	8653	10.4	10.4	745	10.6	10.6
3 Laurea specialistica/magistrale di 2 anni di 2° liv.	550	.7	.7	53	.8	.8
4 Laurea di 3 anni di primo livello o Master di 1° liv.	1203	1.4	1.4	99	1.4	1.4
5 Diploma universitario	658	.8	.8	56	.8	.8
6 Diploma terziario extra-universitario...	406	.5	.5	34	.5	.5
7 Diploma scuola media superiore (4-5 anni)	26066	31.3	31.3	2128	30.3	30.3
8 Diploma scuola media superiore di 2-3 anni	5588	6.7	6.7	651	9.3	9.3
9 Licenza media (o avviamento professionale)	27845	33.4	33.4	2350	33.5	33.5
10 Licenza elementare (o valutazione finale equiv.)	10388	12.5	12.5	788	11.2	11.2
11 Nessun titolo (sa leggere e scrivere)	1095	1.3	1.3	47	.7	.7
12 Nessun titolo (non sa leggere e/o scrivere)	253	.3	.3	8	.1	.1
Total	83375	100.0	100.0	7014	100.0	100.0
Condizione occupazionale	Freq.	%	% valida	Freq.	%	% valida
1 occupato	47225	56.6	56.6	4464	63.6	63.6
2 in cerca di nuova occupazione	7085	8.5	8.5	372	5.3	5.3
3 in cerca di prima occupazione	1223	1.5	1.5	18	.3	.3
4 casalinga	13199	15.8	15.8	909	13.0	13.0
5 studente	721	.9	.9	17	.2	.2
6 inabile al lavoro	816	1.0	1.0	33	.5	.5
7 ritirato dal lavoro	11669	14.0	14.0	1119	16.0	16.0
8 altra condizione	1437	1.7	1.7	82	1.2	1.2
Total	83375	100.0	100.0	7014	100.0	100.0
Posizione occupazionale	Freq.	%	% valida	Freq.	%	% valida
1 dirigente	1627	2.0	2.3	170	2.4	2.6
2 direttivo, quadro	3385	4.1	4.8	368	5.2	5.5
3 impiegato	22390	26.9	31.5	2203	31.4	33.2
4 capo operaio, operaio subalterno e assimilati	25360	30.4	35.7	2470	35.2	37.2
5 apprendista	466	.6	.7	30	.4	.5
6 lavorante a domicilio per conto d'impresa	164	.2	.2	11	.2	.2
7 autonomo come imprenditore	2114	2.5	3.0	167	2.4	2.5
8 libero professionista	3197	3.8	4.5	317	4.5	4.8
9 lavoratore in proprio	7997	9.6	11.3	586	8.4	8.8
10 socio coop. Produz. Beni e/o prestaz. di servizio	313	.4	.4	21	.3	.3
11 coadiuvante	1800	2.2	2.5	133	1.9	2.0
12 collab. Coord. e continuativa (con o senza progetto)	1095	1.3	1.5	86	1.2	1.3
13 prestazione d'opera occasionale	1116	1.3	1.6	72	1.0	1.1
Total	71024	85.2	100.0	6634	94.6	100.0
System	12351	14.8		380	5.4	
	83375	100.0		7014	100.0	

Figura A2. Modello finale ITALIA (output MlwiN).

```
salute_pijk ~ N(XB, Ω)
salute_pijk = β0ijkkkk + 0.065(0.004)Mijk + -0.018(0.000)(eta-gm)ijk + 0.017(0.001)(ascoli-gm)ijk + 0.111(0.009)Borjk +
0.055(0.007)CMIjk + 0.054(0.009)Autjk + -0.008(0.016)Altrojk + -0.085(0.008)disoccupatoijk +
-0.112(0.006)inattivoijk
β0ijk = 3.772(0.008) + v0k + u0jk + e0ijk

[v0k] ~ N(0, Ωv) : Ωv = [0.022(0.001)]
[u0jk] ~ N(0, Ωu) : Ωu = [0.118(0.002)]
[e0ijk] ~ N(0, Ωe) : Ωe = [0.334(0.002)]

-2*loglikelihood(IGLS Deviance) = 169190.337(83375 of 83375 cases in use)
```

Figura A3. Modello finale LOMBARDIA (output MlwiN).

```
salute_pijk ~ N(XB, Ω)
salute_pijk = β0ijkkkk + 0.086(0.015)Mijk + -0.015(0.001)(eta-gm)ijk + 0.014(0.002)(ascoli-gm)ijk + 0.128(0.027)Borjk +
0.035(0.023)CMIjk + 0.052(0.032)Autjk + -0.035(0.087)Altrojk + -0.028(0.034)disoccupatoijk +
-0.098(0.021)inattivoijk
β0ijk = 3.769(0.023) + v0k + u0jk + e0ijk

[v0k] ~ N(0, Ωv) : Ωv = [0.010(0.003)]
[u0jk] ~ N(0, Ωu) : Ωu = [0.101(0.008)]
[e0ijk] ~ N(0, Ωe) : Ωe = [0.315(0.008)]

-2*loglikelihood(IGLS Deviance) = 13660.494(7014 of 7014 cases in use)
```