



RAPPORTI ISTISAN 14|23

ISSN: 1123-3117 (cartaceo) • 2384-8936 (online)

Pt. 1

**Misurare in sanità pubblica:
registri e sorveglianza.**

ASPETTI GENERALI

A cura di G. Costa, C. Cislaghi e S. Salmaso



EPIDEMIOLOGIA
E SANITÀ PUBBLICA

UN SISTEMA PER MONITORARE DISUGUAGLIANZE DI SALUTE E VALUTARE AZIONI DI CONTRASTO¹¹

Giuseppe Costa (a), Nicola Caranci (b), Silvia Stringhini (c), Nicolas Zengarini (d)
(a) Università degli Studi di Torino, Dipartimento di Scienze Cliniche e Biologiche, Orbassano (TO)
(b) Agenzia Sanitaria e Sociale Regionale, Emilia-Romagna, Bologna
(c) Institute of Social and Preventive Medicine, Lausanne University Hospital, Lausanne, Svizzera
(d) Struttura Complessa a Direzione Universitaria, Servizio Sovrazonale di Epidemiologia, ASL TO3 Grugliasco (TO)

Una covariata sociale per l'*equity audit*

La costruzione di dati capaci di informare sul grado di equità nei livelli di salute e di uso dei servizi sanitari dovrebbe accompagnare ogni intervento di contrasto delle disuguaglianze, dal momento della sua progettazione a quello della sua realizzazione e valutazione. Le disuguaglianze di salute siano pervasive e possano nascere in ogni snodo del funzionamento della società. Questo fa sì che ogni strategia di moderazione delle disuguaglianze di salute deve contare sulla collaborazione di tutti gli attori che devono poter fare la loro parte nell'individuare e correggere i meccanismi di generazione che sono sotto la loro responsabilità diretta o di pressione.

Per questo scopo però occorre che ogni attore sia messo nelle condizioni di scrutinare il contesto e le azioni in modo da identificare correggere e valutare le potenziali disuguaglianze di salute che risultano evitabili grazie al suo intervento. Prerequisito indispensabile è dunque che ogni attore abbia a disposizione adeguate informazioni sulla distribuzione sociale dei problemi di salute e delle conseguenze delle azioni di contrasto. Scopo di questo capitolo è di esaminare quanto i sistemi informativi sanitari e statistici in Italia negli anni Duemila siano adeguati a questo bisogno informativo.

Questa esigenza informativa può essere soddisfatta a due livelli di dettaglio, uno più approssimativo utile per scandagliare le implicazioni per le disuguaglianze di salute di ogni azione sociale, e uno più approfondito per identificare con precisione meccanismi e responsabilità.

Il capitolo si limita a dare risposta alla prima esigenza conoscitiva, che potrebbe essere ridotta a questa semplice domanda: come è possibile rendere disponibili in ogni sistema informativo sanitario o statistico adeguate covariate sociali che permettano di descrivere la posizione sociale delle persone e dei contesti in modo tale da poter identificare variazioni sociali nella distribuzione dei fattori di rischio, dei disturbi e delle malattie e dei loro esiti, questo su tutto il territorio nazionale ad ogni livello geografico rilevante per le azioni di contrasto? Solo a questa condizione in Italia ogni attore sarà facilitato nell'esercizio della sua responsabilità di fare *equity audit* delle azioni di sua responsabilità. Questa semplice necessità informativa non è ancora così facilmente esaudibile in Italia negli anni 2010.

¹¹ Questo contributo è il riadattamento del capitolo elaborato da Giuseppe Costa e il Gruppo di Lavoro Interregionale Equità in Salute e Sanità della Commissione Salute delle Regioni dal titolo "Promuovere equità di salute e di sanità in Italia" pubblicato nel volume: Costa G, *et al.* (Ed.). *L'equità nella salute in Italia. Secondo rapporto sulle disuguaglianze sociali in sanità*. Milano: Franco Angeli; 2014.

Fonti disponibili in Italia per monitorare le variazioni sociali nella salute

Nel 2008 il rapporto finale della Commissione sui Determinanti Sociali di Salute OMS proponeva tre raccomandazioni portanti, la terza delle quali è misurare e analizzare le disuguaglianze di salute e verificare l'impatto delle politiche e delle azioni. Un sistema di sorveglianza sull'equità richiede un'adeguata messa a fuoco del ruolo dei determinanti sociali nelle azioni di sanità pubblica ad ogni livello geografico.

Le esperienze realizzate nel contesto europeo per il monitoraggio delle disuguaglianze di vari livelli di evoluzione. Le capacità informative e la disponibilità di covariate sociali vanno di pari passo con l'attenzione dedicata al contrasto delle disuguaglianze, come nel Regno Unito o nei Paesi scandinavi. Anche nel nostro Paese, e tra le Regioni e le Aziende Sanitarie Locali (ASL), sussiste una variabilità in termini di capacità di integrazione e sfruttamento degli apparati informativi sulla salute e sulle covariate sociali.

Di seguito verranno passate in rassegna le principali fonti informative disponibili in Italia per lo scopo di monitoraggio delle variazioni sociali di salute: appartenenti al nuovo sistema statistico sanitario (NSIS) e al sistema statistico nazionale (SISTAN); altre fonti informative accessorie (dati di fonte INPS o INAIL) verranno richiamate di volta in volta nel testo. Il paragrafo si concluderà identificando una serie di punti critici per il miglioramento delle capacità di monitoraggio delle variazioni sociali di salute in Italia, con le relative proposte di sviluppo e correzione.

I sistemi inclusi in NSIS disponibili in Italia costituiscono un'ampia riserva di dati raccolti in modo sistematico e corrente, per scopi prevalentemente amministrativi o gestionali, e riguardanti fenomeni rilevanti per la salute e per l'uso dei servizi. Nonostante la loro missione prevalente di misura dell'uso dei servizi sanitari, essi possono di riflesso offrire una buona stima del bisogno di salute, in tutti i casi in cui il bisogno si traduca in una domanda di servizi, seppur con problemi di completezza e accuratezza che dipendono dallo scopo non statistico ed epidemiologico della fonte informativa. Negli ultimi anni si è gradualmente ampliato lo sfruttamento in termini epidemiologici di questi sistemi di NSIS, con applicazioni che sono andate sotto il nome di epidemiologia dei sistemi informativi. I sistemi compresi in SISTAN hanno il vantaggio di essere più esplicitamente finalizzati alla misura di fenomeni di salute o di uso dei servizi con propositi conoscitivi e quindi di garantire completezza e accuratezza certificate, seppur in modo spesso campionario; lo svantaggio rispetto a NSIS è che per ragioni di efficienza e costo essi non permettono di raggiungere la granularità geografica e temporale garantita dai sistemi di NSIS.

Gli archivi correnti di NSIS e di SISTAN offrono spesso la possibilità di stimare con adeguata accuratezza i numeratori delle misure epidemiologiche (es. numero assoluto di decessi o di ricovero in un anno, per sesso e classe di età). I denominatori (es. popolazioni residenti da cui derivano i decessi o i ricoveri) di solito sono desumibili da altre fonti.

Nella tradizione del nostro Paese, le principali fonti di dati sulle popolazioni sono costituite dalle Anagrafi comunali e dal Censimento della popolazione e delle abitazioni. Le prime sono aggiornate di continuo con i movimenti della popolazione ufficialmente residente, vale a dire con le iscrizioni e cancellazioni per nascita/morte (componente naturale) e per immigrazione/emigrazione (componente migratoria). Essendo affette dalla non perfetta efficienza nella registrazione, i Censimenti svolti ogni 10 anni vengono utilizzati per correggere i contingenti di popolazione e dar conto della popolazione legale.

Le Anagrafi comunali costituiscono tutt'oggi il principale snodo di informazioni sulle popolazioni. I circa 8.100 comuni italiani usano sistemi di registrazione e archiviazione

eterogenei, e molti hanno implementato archivi informatici già da prima degli anni 2000. Con essi si alimentano le basi delle anagrafi delle ASL, che talvolta sono collezionate anche ad un livello regionale. Nell'ambito dell'assistenza sanitaria, poi, gli archivi sono organizzati per assolvere ad altre funzioni, come avviene con l'archiviazione della popolazione presente che sceglie il medico di medicina generale in un comune diverso da quello di residenza oppure con l'eventuale memorizzazione dei contatti con i Sistemi Sanitari Regionali per la richiesta di una prestazione (come può avvenire in un accesso al Pronto Soccorso in una zona lontana da quella di residenza).

L'integrazione dei sistemi per il computo delle misure epidemiologiche può avvenire in una forma più ecologica e trasversale (come accade nel calcolo dei tassi annui di mortalità sulla popolazione media) oppure più analitica e potenzialmente longitudinale (come accade nella costruzione di coorti, seguite nel tempo dove si computa il tempo/persona di ciascun soggetto al denominatore delle misure di occorrenza degli eventi sanitari, che andranno a numeratore). L'approccio su base individuale (più spesso longitudinale) garantisce stime più valide delle variazioni nelle misure epidemiologiche, tutte le volte che l'informazione sulla covariata viene rilevata solo in una delle due fonti o con caratteristiche di accuratezza variabili tra le fonti; questa soluzione però richiede che sia possibile l'integrazione dei dati a livello individuale.

L'integrazione tra archivi anagrafici o di censimento e archivi sanitari di popolazione non è un'operazione fattibile sempre e ovunque, sia per ragioni di riservatezza dei dati personali, sia per limiti del disegno dei sistemi informativi stessi. Esistono Paesi in cui i flussi di dati sono disegnati e predisposti per essere integrati alla fonte, come avviene ad esempio in Svezia, dove esiste un numero identificativo personale applicato in tutte le registrazioni, sia amministrative che statistiche, e dove la protezione della riservatezza viene regolata in sede di accesso e uso dei dati.

Nel caso dello studio delle variazioni sociali nelle misure epidemiologiche occorre disporre di dati sulla posizione sociale, o di altri indicatori di vulnerabilità sociale come l'area di provenienza per gli immigrati (d'ora in poi definiti per semplicità "covariate"). Quasi mai i sistemi informativi che misurano i fenomeni sanitari (NSIS) dispongono di adeguate misure delle covariate sociali, dunque quasi sempre bisogna ricorrere all'integrazione con altre fonti informative che dispongono del dato.

La strada più diretta e coerente per l'acquisizione dei dati sulle covariate sarebbe quella di attingerla dagli archivi che compongono i denominatori e propagarla agli archivi che forniscono i numeratori, in modo da avere un'identica informazione sulla covariata per i casi che verranno contati a numeratore e una maggiore efficienza del sistema, che dovrà rilevare l'informazione una volta sola e potrà propagarla a tutti gli archivi di interesse epidemiologico e di sorveglianza. Per questo scopo la migliore fonte informativa per le covariate è data dal Censimento, che rileva con cadenza decennale le condizioni degli individui (livello d'istruzione, condizione occupazionale e posizione nella professione, stato civile, ecc.), e delle famiglie e abitazioni (composizione familiare, affollamento abitativo, tipologia dell'abitazione, ecc.) (1). Analoga funzione può svolgere l'Anagrafe Tributaria per quanto riguarda le covariate di reddito dichiarato.

Una strada alternativa sarebbe data dalla rilevazione diretta della covariata nella fonte informativa sanitaria di NSIS. In tali casi non sarebbe garantito che l'accuratezza di rilevazione della covariata sia comparabile a quella assicurata dalla fonte utilizzata per il denominatore, ma sono disponibili esempi in cui la soluzione è risultata un accettabile compromesso funzionale, come è avvenuto nel caso delle Schede di Dimissione Ospedaliera (SDO) nelle regioni in cui si rilevi con un adeguato livello di completezza l'ultimo titolo di studio conseguito, o come avviene per tutte le regioni tramite il Certificato di Assistenza la Parto (CedAP, in cui sono disponibili titolo di studio di madre e padre, oltre che condizione occupazionale, stato civile,

ecc.). Invece per quanto riguarda l'origine geografica da molti anni in buona parte dei flussi di NSIS sono rilevate le variabili contenenti il Paese di nascita e la cittadinanza del soggetto che ha avuto un contatto con l'SSN.

Molte informazioni sulla salute e sull'uso dei servizi non sono trattate dai flussi di dati correnti organizzati in NSIS. Per rilevarle sono attive indagini campionarie volte a misurare la salute autopercepita e dichiarata in risposta ad un questionario, insieme ad altre informazioni su comportamenti e stili di vita, oltre che sull'uso dei servizi. Il principale esempio è costituito dal sistema di indagini multiscopo ISTAT sulle condizioni di vita delle famiglie, che comprende una indagine speciale sulla salute e il ricorso ai servizi sanitari, condotta con periodicità circa quinquennale e rappresentativa ad un livello regionale e, dal 2000 anche a livello subregionale. Sullo stesso piano si collocano le indagini condotte per scopo di sorveglianza dalle Regioni e dal CNESPS (Centro Nazionale di Epidemiologia, Sorveglianza e Promozione della Salute dell'Istituto Superiore di Sanità), di cui il Sistema di Sorveglianza PASSI (Progressi delle Aziende Sanitarie per la Salute in Italia) è il principale esempio, in grado di indagare conoscenze, comportamenti e salute di un campione rappresentativo degli adulti di un territorio che comprende circa il 90% della popolazione italiana, con rilevazioni periodiche ravvicinate.

Le indagini campionarie colmano le principali lacune conoscitive sulle variazioni sociali nella salute, rilevando in contemporanea in modo trasversale il dato sulla salute e quello sulla covariata in un campione di popolazione. Per la sua natura trasversale queste indagini possono stimare solo le misure epidemiologiche di prevalenza (e le loro variazioni sociali) e non quelle di incidenza. Per ottenere stime di incidenza occorre seguire nel tempo gli individui del campione per osservare l'insorgenza degli eventi che si desidera studiare. Questo è accaduto innestando un approccio di studio longitudinale sulle indagini speciali multiscopo ISTAT sulla salute, a partire dall'Indagine del 1999-2000 (2). Gli intervistati, e le informazioni ad essi relative nel momento dell'intervista, costituiscono la base per una coorte di soggetti, arruolata e seguita nel tempo tramite gli archivi di dati sanitari correnti (nel caso specifico la mortalità e le dimissioni ospedaliere). Un sistema longitudinale di questo tipo incorpora informazioni sulla popolazione, sulle covariate sociali, sulla salute di partenza e su quali eventi sanitari sperimentano nel tempo, con un'ampia rappresentatività del contesto italiano.

Un altro sistema su base campionaria, che integra dati da flussi correnti è WHIP-Salute (<http://www.dors.it/pag.php?idcm=4433>), una banca dati di storie lavorative individuali relativa a un campione rappresentativo della popolazione attiva e pensionata del settore privato di dimensioni ragguardevoli (7,5% della popolazione sorgente), costruita a partire dagli archivi gestionali dell'INPS, a cui sono state agganciate informazioni di eventi di salute accaduti negli anni successivi (infortuni sul lavoro, malattie professionali, ricoveri, stato in vita) (3).

Naturalmente il limite principale degli studi su base campionaria rispetto alla fonte informativa di NSIS è dato dalla dimensione dello studio campionario, che non potendo avere dimensioni ampie per via dei costi di una rilevazione *ad hoc*, non può assicurare quella granularità geografica fine che gli obiettivi di *equity audit* potrebbero richiedere.

A questa esigenza hanno risposto alcune esperienze metropolitane, le quali, partendo dalla necessità di sottomettere ad *equity audit* le situazioni e le azioni locali, hanno costruito appositi sistemi integrati di osservazione delle variazioni sociali di salute nell'intera popolazione, i cosiddetti studi longitudinali metropolitani, in cui vengono seguite nel tempo le coorti dei residenti censite al censimento di popolazione nel comune, mettendo in relazione gli esiti di salute accaduti con le covariate tratte dal censimento (4).

Agli approcci incentrati sul dato di posizione sociale a livello individuale bisogna aggiungere gli approcci che usano dati di posizione sociale di contesto, cioè riferiti ad un aggregato di taglia appropriata per il contesto che si vuole descrivere. Questa soluzione può essere necessaria sia perché l'informazione di interesse è veramente quella relativa alle caratteristiche sociali del

contesto di vita (es. il fatto di vivere in un quartiere mediamente più povero o più ricco), oppure perché la covariata sociale a livello di aggregato è più facilmente disponibile che non quella a livello individuale (nel caso che nessuna delle soluzioni precedentemente indicate sia disponibile) e quindi bisogna ricorrere alla covariata di aggregato come *proxy* della covariata individuale. Anche nel caso delle covariate a livello di aggregato si pone il problema di integrazione di informazione tra i sistemi informativi; infatti il dato sulla covariata di aggregato non è mai disponibile direttamente nelle fonti informative che forniscono il denominatore e il numeratore delle misure epidemiologiche, occorre dunque creare il dato con la sua fonte informativa diretta (censimento o anagrafe tributaria ad esempio) e poi propagarlo alle fonti informative del denominatore e numeratore della misura epidemiologica (NSIS e SISTAN) attraverso il dato dell'aggregato a cui si riferisce la covariata (regione, provincia, comune, quartiere, sezione di censimento di residenza). Per effettuare questo collegamento tra dati individuali e dati aggregati, basterebbe conoscere per ogni record (es. l'iscritto all'anagrafe sanitaria degli assistiti, il ricoverato, il deceduto, ecc.) il proprio aggregato geografico di residenza, dato che di solito è disponibile per la regione, provincia e comune di residenza, mentre non lo è per la sezione di censimento. Eppure spesso è proprio l'aggregato più piccolo quello che descrive meglio il contesto sociale in cui vive il soggetto. In questi casi si rende necessario acquisire il dato per via indiretta. Le soluzioni possibili, a riguardo, sono principalmente di due tipi: attribuzione tramite anagrafi comunali (con procedure di record *linkage*) oppure tramite georeferenziazione¹² (5, 6).

Quali adeguamenti sarebbero necessari e compatibili?

Priorità di sviluppo

Alla luce dei limiti e delle potenzialità rivelate dalle diverse soluzioni messe in campo nelle esperienze citate, il capitolo proseguirà proponendo un possibile percorso di adeguamento delle capacità dei sistemi informativi italiani di misurare le disuguaglianze di salute in modo capillare su tutto il territorio italiano. In questo percorso l'esperienza dei sistemi anagrafici che registrino o recuperino il dato di posizione sociale e che siano integrati con quelli sanitari via record-*linkage* (RL), viene ritenuta la soluzione più valida ed efficiente, ma sarà trattata solo in relazione a studi locali (del tipo Studi Longitudinali multi-metropolitani) per ragioni di fattibilità. Qualora il contesto normativo e operativo consentisse in modo sufficientemente flessibile il RL tra le fonti su una scala più ampia di quella del singolo comune e della singola anagrafe, allora converrebbe, anche in termini di qualità dei risultati, investire ulteriormente sulle fonti dove la posizione sociale viene rilevata meglio (anagrafe e censimento), per propagarla a quelle con informazioni sanitarie.

In mancanza/attesa di questa soluzione, l'alternativa è che si rilevi appositamente il dato sulla posizione sociale nella fonte sanitaria, come già accade nella SDO con il livello

¹² Quest'ultima tecnica ha come prerequisito la conoscenza degli indirizzi per definire la posizione geografica dei casi e la disponibilità di strumenti per georeferire i dati, vale a dire rintracciarli in uno stradario in cui siano presenti gli stessi indirizzi, reperire le coordinate di latitudine e longitudine dei punti e collocare i punti in una mappa. In tal modo è possibile proiettare i punti ottenuti su una mappa delle sezioni di censimento e cercarne l'intersezione con i poligoni che disegnano i confini di ogni piccola area. Rintracciate le intersezioni, si arriva a disporre dell'informazione sull'area (sezione di censimento), di cui si conosce già la posizione sociale. In aree non urbane può risultare sufficiente usare dati relativi al comune, informazione questa generalmente presente in tutti gli archivi di dati correnti.

d'istruzione seppure con severi problemi di completezza e accuratezza, o nel CedAP con i dati di posizione sociale dei genitori. Quando nemmeno questa soluzione fosse disponibile, rimane solo l'applicazione alla fonte sanitaria del dato sulla covariata sociale dell'aggregato di residenza ricorrendo a sistemi informativi geografici e tecniche di georeferenziazione per la contemporanea analisi dell'informazione su base spaziale.

Naturalmente nel caso che non sia indispensabile monitorare le disuguaglianze sociali di salute a livello capillare nello spazio e nel tempo, ma ci si possa accontentare di una valutazione temporale periodica e su scala nazionale o al più regionale, allora il ricorso alle indagini campionarie trasversali e longitudinali presenta il massimo di validità ed efficienza e su questo versante dell'offerta statistica la situazione italiana non ha bisogno di essere ulteriormente migliorata.

Cruciale in questo percorso di adeguamento italiano alla disponibilità di covariate sociali è assicurare la completezza, l'affidabilità delle variabili, il coordinamento e la collezione delle anagrafi a cui riferire i dati sanitari, perché esse sono alla base di differenti modalità con cui vengono organizzati e integrati i dati NSIS. Le ASL e gli operatori che gestiscono la compilazione e la manutenzione dei sistemi informativi hanno un ruolo centrale, che può essere efficientemente affiancato da sistemi anagrafici regionali o nazionali che, nel rispetto dei criteri dettati dalla normativa sui dati personali, permettano un'univoca identificazione dei soggetti. Tramite gli archivi di dati sulla popolazione assistibile si ottiene non solo una stima dei denominatori delle misure epidemiologiche, ma si possono caratterizzare i soggetti con tutte le covariate sociali individuali ricavabili da fonti informative integrabili con l'anagrafe, e con i dati per la localizzazione di assistiti e di punti di erogazione dell'assistenza che servono per l'attribuzione di covariate sociali di aggregato.

Di seguito verranno valutate le implicazioni di queste priorità di sviluppo.

Quando la covariata sociale è tratta da dati aggregati

L'uso di dati geografici è la via spesso più rapida o l'unica possibile per caratterizzare i soggetti in studio, sia tramite gli attributi del contesto più o meno ampio in cui essi vivono, sia per conferire loro la posizione sociale dell'aggregato per approssimarne la posizione individuale. Tra i dati di contesto vi sono quelli prettamente areali, come l'altitudine o il grado di urbanizzazione, e quelli che sintetizzano misure composite (es. benessere: www.misuredelbenessere.it). Allo stesso modo, tra i dati che approssimano la condizione individuale si ritrovano misure semplici (es. reddito mediano familiare equivalente per sezione di censimento) o misure multidimensionali, come l'indice di deprivazione basato sul censimento, che saranno tendenzialmente migliori *proxy* della condizione individuale quanto più piccola sarà l'area a cui si riferiscono. Le sezioni di censimento sono l'aggregato amministrativo più piccolo disponibile in Italia, con una media di meno di 200 residenti (e una variabilità molto accentuata).

La fattibilità di tale approccio, che sarebbe rapido e poco dispendioso, è legata all'effettiva disponibilità dei dati (pubblici ma talvolta difficilmente accessibili, come nel caso dei redditi). Inoltre, nel caso si vogliano usare indicatori a livello di aggregati inferiori al comune (il cui dato è rilevato in tutti i sistemi informativi, in relazione alla residenza degli assistiti), bisogna localizzare gli indirizzi (il più delle volte di residenza) dei soggetti. Per compiere tale operazione è molto rilevante disporre di una sua registrazione di buona qualità. Per rintracciarne le coordinate nello spazio sono disponibili applicativi in Internet, che hanno limiti nell'approssimazione delle coordinate (soprattutto in ambito non urbano) e un tetto nel numero di record nel processare massivamente archivi di indirizzi. Il processo che porta alla localizzazione dei dati, detto anche georeferenziazione, è composto da due fasi: ricerca

dell'indirizzo in un indirizzario "standard", tramite una normalizzazione delle stringhe che lo descrivono (per trasformarlo in una dizione univoca), e poi nell'attribuzione delle coordinate spaziali.

In alcune regioni, come l'Emilia-Romagna e la Toscana, sono disponibili sistemi per la normalizzazione e attribuzione delle coordinate appositamente sviluppati per i relativi stradari e per anni recenti. Inoltre per l'intero Paese è in corso un'attività per la costruzione di un archivio nazionale "delle strade e dei numeri civici geocodificato alle sezioni di censimento aggiornate dall'ISTAT al 2010 (7). Una simile attività è stata inserita nel Piano Statistico Nazionale 2011-2013 (aggiornamento 2013; IST-02495) (8). Inoltre l'ISTAT ha annunciato nuovi "strumenti GIS per aggiornare le basi territoriali per i censimenti" e il rilascio dei limiti amministrativi di regioni, province e comuni, aggiornati al 1° gennaio 2011. Sono infatti disponibili i dati geografici delle basi territoriali dal Censimento 1991 al 2011.

L'alternativa alla georeferenziazione degli indirizzi e alla loro caratterizzazione con il dato della sezione di censimento è data dall'acquisizione diretta della sezione di censimento tramite il collegamento con l'anagrafe comunale, come avvenuto nel caso delle esperienze dell'attribuzione del reddito ai dati dei ricoveri ospedalieri già citato. La soluzione in questo caso richiederebbe che la sezione di censimento entri a far parte correntemente degli archivi anagrafici a supporto di quelli sanitari, ad esempio nelle anagrafi sanitarie, nello stesso momento in cui queste acquisiscono i dati dai Comuni e li aggiornano con i relativi movimenti della popolazione.

Quando la covariata sociale è presente in archivio NSIS

Qualora sia necessario o preferibile disporre di informazioni individuali sulla condizione socio-demografica (come nel caso della provenienza geografica da altro Paese) e non si disponga di queste informazioni da fonte anagrafica (e dunque anche sul denominatore), allora sarà particolarmente utile avere a disposizione negli archivi sanitari variabili appositamente rilevate.

I flussi di dati regolamentati a livello nazionale e con una raccolta sistematica e standardizzata, come quelli definiti dal Ministero della Salute, rilevano in modo diffuso le informazioni sulla provenienza geografica e meno spesso le informazioni sulla posizione sociale. Alcuni flussi costituiscono un'eccezione; a titolo esemplificativo della resa di questa soluzione si dà sinteticamente conto di due di questi (SDO, CedAP) e un accenno ad altri due (FAR - Assistenza Residenziale e semi-residenziale per anziani, Salute Mentale Adulti).

Il contenuto del flusso di dati sulle SDO è definito dal Ministero della Sanità tramite il Decreto 27 ottobre 2000 (9). L'articolo 1, al punto 6 bis, stabilisce che tra le informazioni da rilevare ci sia anche il livello d'istruzione, codificato in 5 modalità. Tuttavia la compilazione dell'informazione è facoltativa e incompleta; nel 2011 a livello nazionale la distribuzione presenta ancora delle criticità, con un 35% di dati non validi.

In qualche caso non è possibile chiedere l'informazione al paziente (oppure il contesto non permette una rilevazione accurata, come si può ipotizzare se il ricovero avviene in regime d'urgenza) o la richiesta non è utile e pertinente, perché il soggetto è in età in cui non ha ancora maturato nessun titolo. Bisogna però tener in conto che in alcune regioni italiane il livello d'istruzione è stato registrato con migliore completezza (come in Piemonte a seguito dell'aggiornamento della normativa con un 14,2% di dati mancanti nel 2000) (10). Inoltre, data l'alta frequenza di ricoveri, soprattutto in alcune fasce di popolazione, la ricerca del dato sul titolo di studio su almeno uno dei ricoveri dello stesso soggetto permetterebbe di migliorare questa percentuale di insuccesso, caratterizzando così la posizione sociale di un alto numero di eventi. La qualità è variabile e può dipendere dal tipo di ricovero; l'affidabilità è ipoteticamente

migliore in ricoveri non acuti. Per migliorare la completezza e la qualità della compilazione di questa variabile nel record della SDO occorre che sia la programmazione regionale sia quella aziendale considerino al meglio la rilevanza di tale informazione per l'*equity audit* di questo livello di assistenza, attraverso opportune formule di incentivazione e sanzione.

Il caso del CedAP è un caso più fortunato per gli scopi di rilevazione diretta del dato sulla covariata sociale. La rilevazione del dato avviene in una circostanza quasi sempre più favorevole rispetto alla compilazione della SDO; il questionario viene somministrato alla donna in seguito al parto, in condizioni di maggiore tranquillità rispetto a quella di un soggetto ospedalizzato per ricoveri acuti.

Dall'inizio degli anni 2000, il flusso CedAP (11) è la fonte più ricca a livello nazionale di informazioni sanitarie, epidemiologiche e socio-demografiche relative all'evento nascita. Sono qui acquisite informazioni socio-demografiche sui genitori: età al parto, residenza, titolo di studio e condizione professionale di entrambi i genitori, oltre che il loro stato civile e la cittadinanza.

La compresenza delle informazioni su gravidanza, stato del neonato e delle informazioni sulla famiglia di provenienza, con livelli di completezza delle variabili che le rilevano tendenzialmente superiori al 90% (12), consente di stimare misure epidemiologiche di occorrenza e associazione relativamente alla nascita e alle condizioni influenti. In tale circostanza, si dispone quindi contemporaneamente del dato sul denominatore (parti) e sul numeratore (esiti della gravidanza e del parto) delle misure epidemiologiche. Inoltre si possono impiantare, sulla base di tali dati, coorti di nascita da seguire nel tempo tramite archivi NSIS. L'operazione è fattibile senza perdita di informazioni e senza dispendio di tempo/lavoro a condizione che siano facilmente associabili le registrazioni del Certificate con quelle del neonato, che se è a sua volta identificato univocamente negli archivi NSIS può essere incluso in un'osservazione prospettiva. È il caso di quanto sperimentato sullo studio dell'ospedalizzazione e della mortalità infantili in Emilia-Romagna (13).

Per quanto riguarda gli altri archivi che contengono almeno un'informazione sulla posizione sociale individuale dei soggetti che rientrano nella rilevazione, si segnalano altri due flussi informativi nazionali:

- flusso informativo regionale FAR (Assistenza Residenziale e semi-residenziale per anziani), attivato anche a livello nazionale dal 2010 (14). Rileva, tra le altre variabili, il titolo di studio in 6 modalità;
- salute mentale adulti, oggetto di un Accordo della Conferenza Stato Regioni (stipulato il 11/10/2001 su "Il Sistema Informativo Nazionale per la Salute Mentale"), e che per ogni paziente prevede di rilevare il titolo di studio e la condizione occupazionale.

A riguardo della possibilità di mettere a frutto l'informazione sulla provenienza da altro Paese, particolarmente rilevante negli ultimi anni data la rapida crescita della popolazione immigrata e l'emergere di altri bisogni di salute, gli archivi NSIS contengono le informazioni necessarie sui numeratori di possibili indicatori di salute e hanno raggiunto negli anni un buon livello di qualità e completezza¹³, pur persistendo qualche differenza tra le diverse regioni italiane. Sarebbe tuttavia utile disporre di tale informazione accurata anche per i denominatori e per dati non sanitari, in modo da poter stimare correttamente gli indicatori per la più ampia definizione del profilo di salute degli immigrati, e da poterli calcolare anche per quelli non regolari.

13 Ad esempio, la percentuale di SDO degli stranieri con campo cittadinanza vuoto o errato, a livello nazionale, è costantemente dello 0,1% negli anni 2007-2009. A livello nazionale, nel 2008, i parti rilevati con il CedAP sono pari al 96,6% di quelli rilevati con le SDO e il numero di nati vivi pari al 98,1% di quelli registrati presso le anagrafi comunali. Non è stata ancora studiata, a livello nazionale, la qualità della variabile cittadinanza.

Quando la covariata sociale è tratta da archivi di dati non NSIS

Le fonti sanitarie di NSIS registrano i contatti con e le prestazioni del Servizio Sanitario Nazionale e acquisiscono così informazioni con un approccio “oggettivo” su tipo e data dell’intervento e, a volte, diagnosi. Ma non sempre i bisogni di salute si traducono in domanda e fruizione di prestazioni; inoltre ci sono disturbi, comportamenti e condizioni influenti, oltre che aspetti “soggettivi” della salute, che non vengono rilevati con i dati delle fonti sanitarie su tutta la popolazione. Le indagini campionarie che interrogano direttamente l’interessato (o suoi familiari) costituiscono un’originale fonte di informazioni che da sola rileva dati su eventi sanitari e di tipo sociale.

Per valutare contemporaneamente il contesto familiare, sociale e culturale con quanto riferito su percezione della salute, uso dei servizi, comportamenti e stili di vita, come precedentemente riportato esistono in Italia due indagini campionarie condotte periodicamente e con sistematicità: la Multiscopo su Condizioni di salute e ricorso ai servizi sanitari e il Sistema di sorveglianza PASSI. Si basano su interviste e sono condotte dall’ISTAT e dalle Aziende Sanitarie.

Le indagini campionarie hanno costi contenuti e se il campionamento è corretto i risultati sono generalizzabili. Essendo trasversali, però, non sono in grado di valutare la direzione della relazione causale nelle disuguaglianze che osservano, se cioè la relazione osservata tra disoccupazione e peggior salute sia dovuta all’effetto nocivo della disoccupazione o al fatto che la salute compromette il successo nella ricerca del lavoro. Inoltre esse non consentono di rilevare eventi rari o di breve durata; e i loro risultati possono essere distorti a causa di problemi di incompletezza nella rispondenza o di accuratezza dell’informazione rilevata dai questionari. Alcuni di questi limiti sono superabili con la creazione di coorti longitudinali sulla base del campione intervistato. Grazie all’armonizzazione delle variabili per collegare i dati, si sono sviluppati studi longitudinali basati su coorti storiche, ottenute dall’integrazione tra fonti informative campionarie e fonti di dati amministrativi e sanitari. Il punto di forza di questi sistemi è che rilevano, attraverso fonti sanitarie amministrative e statistiche, gli eventi di salute che sono capitati nel tempo agli individui che sono stati arruolati in coorti per indagini precedenti, attraverso cui si conoscono le caratteristiche sociali della persona. Il più significativo di questi studi per gli scopi di monitoraggio delle disuguaglianze di salute è lo Studio Longitudinale Italiano (SLI), inizialmente basato sul *record-linkage* tra il campione dell’indagine Multiscopo sulla salute (ISTAT 1999-2000) e gli archivi delle schede di morte e dei ricoveri 1999-2007. La coorte è arruolata attraverso l’indagine multiscopo, dove la posizione sociale della persona e della famiglia è ben misurata, oltre ai dati di partenza su salute, stili di vita e di ricorso ai servizi. Essa poi viene seguita per quanto riguarda la morbosità attraverso la ricerca delle cause di ricovero e delle cause di morte negli anni successivi. In questo modo è possibile studiare l’effetto delle condizioni sociali della persona sulla morbosità e l’uso dei servizi. Tale studio è riconosciuto tra quelli SISTAN, in quanto inserito nel Piano Statistico nazionale (IST-02269) (15), e prevede di connettere gli archivi ISTAT delle indagini multiscopo sulla salute di più edizioni successive a quella del 1999-2000 (2004-2005, 2012-2013 e quella speciale sugli immigrati stranieri) con schede di morte e di ricovero, e anche con sviluppo sui dati di ricorso alla assistenza specialistica e a quella farmaceutica.

Un secondo sistema di monitoraggio longitudinale su base campionaria, utile per lo studio delle disuguaglianze sociali di salute che nascono nel mondo del lavoro e della previdenza, è stato realizzato per l’analisi dei destini di salute delle carriere lavorative del settore privato tramite un campione estratto dall’archivio nazionale dell’INPS, con le relative ricostruzioni dal 1985 al 2005 (*Work Histories Italian Panel, WHIP*. In PSN: SAL-00053) (16). Ai soggetti campionati sono stati collegati individualmente le registrazioni di fonte INAIL degli infortuni

sul lavoro accaduti tra il 1994 al 2005, con un *record-linkage* basato sul codice fiscale criptato, ottenendo un archivio denominato WHIP-Salute.

I principali limiti per la misura delle disuguaglianze di salute di questi studi longitudinali su base campionaria e su scala nazionale sono la natura campionaria (con le conseguenti limitazioni di potenza statistica nel caso di SLI) e la restrizione dell'osservazione alle poche fonti informative statistiche e sanitarie integrabili a livello nazionale. Per ovviare a tali limiti, si sono attivati studi longitudinali basati sull'integrazione di archivi su base cittadina, ad iniziare dal Censimento di popolazione. A livello di un singolo comune si può disporre dei dati dell'intera popolazione e si può integrare il dato censuario con quelli delle fonti sanitarie di NSIS disponibili a livello locale o regionale. Nel caso di comuni di grandi dimensioni, si superano così i limiti della potenza statistica. Questo tipo di esperienza si sta diffondendo in diversi capoluoghi italiani, e nei progetti avviati già da alcuni anni, stanno assicurando una visione longitudinale retrospettiva molto profonda, cosa che permette uno studio più approfondito delle dinamiche delle disuguaglianze sociali di salute. Il più "storico" di questi studi è quello della città di Torino, che parte dal censimento 1971; ad esso è seguito lo studio Toscana (dal 1981 con Livorno e poi con Firenze e Prato) e quello di Reggio Emilia (sul censimento 1991). Si è poi costituita una rete nazionale degli studi longitudinali, definita anche tramite una scheda del Programma Statistico Nazionale (PSN 2012-2013; scheda EMR-00018), che include altri tre Studi: lo Studio emiliano, con Bologna, Modena e Reggio Emilia, lo Studio di Roma e quello di Venezia. Un ulteriore studio è in corso di attivazione (Palermo e Catania). Il primo censimento da cui partono le coorti di questi ultimi Studi Longitudinali è quello 2001.

La fattibilità dell'integrazione con dati di censimento e la creazione di coorti censuarie garantisce ampie possibilità di indagine su buona parte delle dimensioni delle disuguaglianze di salute. Offre un modello per l'integrazione di dati sull'universo degli individui, con l'opportunità di studiare intere popolazioni. Tuttavia l'approccio è affetto da alcune limitazioni: le covariate sociali sono aggiornate solo in corrispondenza dei censimenti della popolazione, che si replicano ogni 10 anni, con quello 2011 che ne rileva una buona parte e che in futuro dovrebbe essere continuo a rotazione solo su frazioni di popolazione (16). Inoltre è essenziale che i Comuni coinvolti abbiano un'anagrafe storicizzata, con la registrazione di variabili chiave che consentano il collegamento con i censimenti; inoltre è richiesta la necessaria attività per il corretto abbinamento degli archivi.

Gli archivi di popolazione, arricchiti con dati registrati per finalità non strettamente anagrafiche, sono lo snodo che concorrerebbe maggiormente ad ampliare la capacità di descrivere i denominatori delle misure rilevanti in sanità e, una volta centralizzati, porterebbero ad un guadagno di efficienza.

Bibliografia

1. Caranci N, Costa C, Pacelli B, Fano V. Censimento della popolazione: informazione cruciale per valutare il rapporto tra società e salute. *SISmagazine*, 12 aprile 2011. Disponibile all'indirizzo: http://old.sis-statistica.org/magazine/spip.php?article198&var_recherche=caranci; ultima consultazione 3/11/14.
2. Marinacci C, Grippo F, Pappagallo M, Sebastiani G, Demaria M, Vittori P, Caranci N, Costa G. Social inequalities in total and cause-specific mortality of a sample of the Italian population, from 1999 to 2007. *Eur J Public Health* 2013;23(4):582-7.
3. Bena A, Giraudo M. Lavoro temporaneo e salute: un'analisi del rischio infortunistico per durata del contratto di lavoro. *Epidemiol Prev* 2013;37(1):29-34.

4. Costa G, Demaria M. Un sistema longitudinale di sorveglianza della mortalità secondo le caratteristiche socio-economiche, come rilevate ai censimenti di popolazione: descrizione e documentazione del sistema. *Epidemiologia e Prevenzione* 1988;36:37-47.
5. Sabatti V, Crescenzi F. *Georeferenziazione dei dati e integrazione delle fonti locali*. Roma: Istituto Nazionale di Statistica, 1999.
6. Petrucci A, Bocci C, Borgoni R, Civardi M, Salvati N, Salvini S, Vignoli D. *Indagine sulla georeferenziazione dei dati nella statistica ufficiale*. Roma: CoGIS, 2009.
7. Istituto Nazionale di Statistica. *Nuovo archivio nazionale per strade e numeri civici*. Roma: ISTAT; 2010. Disponibile all'indirizzo: http://www.istat.it/salastampa/comunicati/non_calendario/20100706_01/; ultima consultazione 18/12/14.
8. Istituto Nazionale di Statistica. Programma statistico nazionale 2011-2013. Aggiornamento 2013. Volume 1: Evoluzione dell'informazione statistica. Roma: ISTAT; 2013. Disponibile all'indirizzo: http://www.sistan.it/fileadmin/Repository/Home/PSN/Programma_statistico_nazionale/PSN_2011-2013_aggiornamento_2013/Volumi/Vol.1_Psn_Agg_2013.pdf; ultima consultazione 18/12/14.
9. Italia. Decreto Ministero della Sanità 27 ottobre 2001, n. 380. Regolamento per l'aggiornamento della disciplina del flusso informativo sui dimessi dagli istituti di ricovero pubblici e privati. *Gazzetta Ufficiale - Serie Generale n. 295, 19 dicembre 2000*.
10. Costa G, Spadea T. Le fonti correnti dei dati per la misura delle disuguaglianze di salute: adeguatezza, accessibilità, integrazione. In: *Convegno Associazione Italiana Epidemiologia: Metodi e strumenti per la misura delle disuguaglianze*. Roma, 15-16 maggio 2008. disponibile all'indirizzo: http://www.epicentro.iss.it/focus/aie/Aie08_Costa.asp; ultima consultazione 18/12/14.
11. Italia. Decreto Ministero della Sanità 16 luglio 2001, n. 349. Modificazioni al certificato di assistenza al parto, per la rilevazione dei dati di sanità pubblica e statistici di base relativi agli eventi di nascita, alla natimortalità ed ai nati affetti da malformazioni. *Gazzetta Ufficiale - Serie Generale n. 218, 19 settembre 2001*
12. Giunta Regionale, Direzione Generale Sanità e Politiche Sociali. *La nascita in Emilia-Romagna. 5° Rapporto sui dati del Certificato di Assistenza al Parto (CedAP) - Anno 2007*. Regione Emilia-Romagna, 2008. Disponibile all'indirizzo: http://salute.regione.emilia-romagna.it/siseps/sanita/cedap/files/rapporto_2007.pdf; ultima consultazione 18/12/14.
13. Baronciani D, Battaglia S, Caranci N, Lupi C. *Ricoveri e mortalità infantile: il ruolo delle disuguaglianze. Una analisi effettuata incrociando dati correnti (CEDAP, SDO e REM) degli anni 2007-2011*. Regione Emilia-Romagna, 2012. Disponibile all'indirizzo: <http://assr.regione.emilia-romagna.it/it/servizi/pubblicazioni/rapporti-documenti/ricoveri-e-mortalita-infantile-il-ruolo-delle-disuguaglianze.-una-analisi-effettuata-incrociando-dati-correndi-cedap-sdo-e-rem-degli-anni-2007-2011>; ultima consultazione: 18/12/14
14. Italia. Decreto Ministero del lavoro, della salute e delle politiche sociali 17 dicembre 2008. Istituzione della banca dati finalizzata alla rilevazione delle prestazioni residenziali e semiresidenziali. *Gazzetta Ufficiale - Serie Generale n. 6, 9 gennaio 2009*.
15. Istituto Nazionale di Statistica. Programma statistico nazionale 2011-2013. Aggiornamento 2013. Volume 2: dati personali. Roma: ISTAT; 2013. Disponibile all'indirizzo: http://www.sistan.it/fileadmin/Repository/Home/PSN/Programma_statistico_nazionale/PSN_2011-2013_aggiornamento_2013/Volumi/Vol.2_Psn_Agg_2013.pdf; ultima consultazione 18/12/14.
16. Istituto nazionale di statistica. Programma statistico nazionale 2011-2013. Aggiornamento 2012-2013. Volume 3: dati personali. Roma: ISTAT; 2013. Disponibile all'indirizzo: http://www.sistan.it/fileadmin/Repository/Home/PSN/Programma_statistico_nazionale/PSN_2011-1013_aggiornamento_2012-2013/Volumi/Volume_3_Psn_2012.pdf; ultima consultazione 18/12/14.