

Big data, un'opportunità da non perdere

Riccardo Giampaoli¹, Giuseppe Quintaliani²

¹SC Chirurgia Vascolare, Azienda Ospedale Perugia, Perugia

²Responsabile Governo Clinico della SIN, Roma

Big data, big value, huge opportunity

The constant computerization process, in a liquid and continually changing society, characterizes all of our activities, bringing out the need to develop new methodologies that will help the institutions to maintain the social status.

These thoughts are becoming more significant in fields as important as that of National Health, which is incessantly hit by the spending review cuts and that constantly tries to keep in order the balance sheet along with a high level of care.

In this context the big data analysis has a key role and is seen as a great antidote to the approximation that has distinguished our supervising managers by highlighting the most trustworthy correlations that can be used in all the different settings of the health system.

In our country the big data have already been applied in National Health institutions with promising results and have challenged the administrations, whose principal goals consist in maintaining the highest levels of care and correct balance sheets, that will have to recognize the opportunities and advantages that the big data analysis can offer.

Keywords: Big data, Big data analysis, Health and big data



Riccardo Giampaoli

Introduzione

George Orwell pubblica, nel 1984, il suo romanzo "Grande fratello", immaginando una società sotto totale controllo. Da allora, molte cose sono cambiate e, soprattutto, almeno in un certo senso, l'interconnessione digitale ha quasi avverato quanto scritto da Orwell, nel senso che tutto è tracciabile anche se non controllabile.

Le tecnologie legate all'*Information Computer Technology* (ICT) hanno attenuato in noi il rigido controllo sui dati personali che fino a pochi anni fa conservavamo gelosamente (uso delle *mail*, uso delle ricerche e consultazione di *Internet* anche dal telefonino). Va tenuto presente che ogni comportamento *online* realizzato con il

supporto di uno strumento dotato di indirizzo IP lascia traccia di un aspetto della nostra vita: interessi, gusti e preferenze, interazioni sociali, posizione geografica, situazione economico-finanziaria, comportamenti di acquisto e così via. Allo stesso modo, l'*Internet of Things* (IoT) permette che le informazioni raccolte da sensori inseriti in oggetti di uso quotidiano vengano trasformate in dati e condivise in rete per essere analizzate e monitorate. È, così, possibile registrare e archiviare dati su aspetti diversi e inimmaginabili della vita quotidiana.

In altre parole, ogni volta che si usa un *computer*, è possibile tracciare cosa si è fatto, cosa si è cercato e che soluzioni abbiamo adottato. Il sistema non è in grado di dire chi ha compiuto una certa operazione ma sicuramente è in grado di sapere quanti hanno cercato una certa cosa ed è in grado di riproporre al *computer* da cui è stata fatta la ricerca offerte e proposte. Provate a fare una ricerca su *Google* di un *computer* e avrete offerte di *computer* anche quando andate su *Facebook*, quando navigherete di nuovo anche su siti di turismo e così via. *Google* è in grado di dire in ogni momento quante persone hanno acquistato *computer*, quale e/o quali vacanze sono state scelte dagli italiani o solo dagli abitanti di Perugia e così via. Insomma, una massa enorme di dati che viene chiamata *big data*, che, opportunamente elaborata, interpretata e analizzata è in grado di dare informazioni macro su aspetti inimmaginabili della società.

Si parla di analisi dei *big data* quando si ha un *database* talmente grande da richiedere strumenti informatici adeguati

Accepted: April 19, 2016

Published online: May 12, 2016

Indirizzo per la corrispondenza:

Dr. Riccardo Giampaoli
SC Chirurgia Vascolare
Ospedale Santa Maria Misericordia
Piazzale Menghini 1
06156 Perugia
g.quintaliani@yahoo.it

per estrapolare, gestire e processare informazioni entro un tempo ragionevole (1).

L'elaborazione dei *big data* richiede, quindi, elevate capacità di calcolo, tecnologie e risorse che vanno ben oltre i sistemi convenzionali di gestione e immagazzinamento dei dati usati in passato (2), ma che oggi, grazie all'evoluzione tecnologica, siamo in grado di processare con costi minimi e relativa semplicità.

La prima cosa da dire è che il valore autentico dei dati non è quello che emerge a prima vista, ma può essere paragonato a un *iceberg* che si sposta lentamente nell'oceano, dal quale emerge solo la punta, mentre tutto il resto resta nascosto sotto la superficie. Il vero potenziale dell'analisi *big data* emergerà nel momento in cui i dati prodotti non verranno più considerati un patrimonio statico la cui utilità si esaurisce una volta realizzato lo scopo per cui sono stati raccolti, ma quando, attraverso l'analisi incrociata di più *dataset*, ovvero di determinate variabili all'interno di un *database*, magari creati originariamente per tutt'altro scopo, gli stessi potranno, allora, trasformarsi in una possibile "miniera d'oro" di informazioni. Le correlazioni che ne emergeranno saranno in grado di trasformare le banche dati, che le nostre amministrazioni già possiedono, in delle vere e proprie fonti conoscitive dalle quali ricavare indicatori incredibilmente attendibili.

Big data in Sanità

Anche in sanità, il numero di dati prodotti in digitale e, quindi, ottenibili è enorme; lo dimostrano il crescente numero di ricerche mediche effettuate anche sul web nonché il costante processo dell'informatizzazione che sta caratterizzando l'evoluzione delle nostre aziende sanitarie (3).

Per quanto riguarda il nostro paese, il D.L. 179/2012, "Fascicolo Sanitario Elettronico", pone l'attenzione sull'elaborazione dei "big data sanitari", termine che definisce l'insieme di dati e documenti digitali di tipo sanitario e socio-sanitario generati da eventi clinici presenti e trascorsi, riguardanti l'assistito.

Va detto che, se c'è un ambito importante nel quale i *big data* possono essere molto efficaci, questo è proprio quello della sanità pubblica. Proprio attraverso questo nuovo tipo di analisi sarà possibile comprendere le malattie, potendo incrociare l'esito di analisi cliniche con dati ambientali e stili di vita, ottenendo, così, correlazioni in grado di portare consistenti vantaggi in termini di prevenzione.

In questo caso, l'analisi dei *big data* potrà essere uno strumento per la previsione e la riduzione del rischio ma anche per la corretta allocazione delle risorse in ambito sanitario. Alla luce di ciò, è plausibile sostenere che, attualmente, il valore aggiunto di un'azienda che possiede delle banche dati, non risiede unicamente nelle modalità innovative di raccolta delle informazioni o nei singoli dati che detiene e nemmeno nella propria competenza tecnica, ma nelle idee e nel saper cogliere le opportunità che una mentalità orientata all'analisi dei *big data* può offrire.

Un esempio concreto a livello internazionale di tale opportunità è quello del *MedStar Hospital Center* di *Washington* (U.S.A.) che si avvale della collaborazione di *Microsoft Research*, il comparto ricerca e sviluppo *software* di *Microsoft*: quest'ultimo ha analizzato, attraverso il *software Amalga*, parecchi anni di *report* sanitari resi anonimi, come dati demografici, esami clinici, diagnosi e terapie, alla ricerca di soluzioni che permettessero la riduzione dei ricoveri e delle infezioni ospedaliere. L'uso di questa tecnica ha evidenziato una serie di fattori che incentivavano l'aumento di nuovi ricoveri dei pazienti a distanza di un mese dalla loro dimissione, facendo, inoltre, emergere una forte correlazione tra tale dinamica e le diagnosi iniziali in cui i pazienti, oltre alla patologia per la quale venivano ricoverati, presentavano anche problematiche psichiche di vario tipo, come depressione o semplice ansia.

La correlazione emersa dall'analisi dei dati ha evidenziato come la compresenza di tali disturbi determinasse con altissima frequenza nuovi ricoveri di questi pazienti nel mese successivo alla loro dimissione, particolarità che, tra l'altro, generava un notevole aumento della spesa sanitaria.

Ciò ha permesso alla direzione del *MedStar Hospital* di elaborare una strategia che prevedesse interventi di sostegno psicologico post-dimissione, finalizzati al miglioramento del benessere psico-fisico dei pazienti che, pur ricoverati per tutt'altra patologia, presentavano anche questi disturbi. Grazie alle informazioni ricavate si è arrivati a ridurre in maniera significativa il numero di nuovi ricoveri ospedalieri e a produrre un conseguente e sostanziale risparmio economico per la stessa azienda.

Per quanto riguarda il panorama italiano, va invece detto che, nonostante in molte regioni si sia proceduto a un primo processo di informatizzazione delle amministrazioni, al momento l'opportunità di sfruttare le potenzialità dell'analisi dei *big data* in sanità non è stata ancora colta pienamente.

Dobbiamo ricordare che, comunque, la quantità di dati prodotta dai sistemi sanitari è enorme, anche se non tutte le realtà sono allineate. In altre parole, molte realtà e organizzazioni sanitarie producono grandi quantità di dati sanitari che, se dapprima erano, e in molti casi rimangono, slegate tra loro, vengono comunque automaticamente e indipendentemente prodotte e sarebbero abbastanza facilmente utilizzabili nell'ottica di *big data*. Infatti, abbiamo il sistema DRG basato sulla SDO (Scheda di dimissione ospedaliera obbligatoria per legge) che contiene i codici ICD9 ma anche il reparto di dimissione, oltre a giorni di degenza e anagrafica paziente, che ormai è ubiquitario e che spesso è usato per ricerche anche da elementi e organizzazioni blasonate e che è il cuore del sistema AGENAS di sorveglianza sanitaria. Ma l'elenco dei *database* sanitari è sterminato: laboratorio e radiologia (che ormai ubiquitariamente sono informatizzati, anche se spesso non raggiungibili in rete), lettere di dimissione, esenzioni di patologia, archivio attività ambulatoriale, spesa farmaceutica, *Istat* mortalità, residenza

(permette di geolocalizzare le patologie) e medico di base e l'elenco è in continuo aggiornamento. Tutti questi *database* sono collocati in sistemi diversi che spesso non colloquiano tra loro e non hanno punti di contatto, ma vengono creati ogni giorno e l'unica cosa da fare sarebbe metterli in contatto e condividerli con sistemi potenti di trasmigrazione e analisi che abbiano un codice unico attribuibile al singolo paziente attraverso la sua trasmigrazione in codice unico secondo algoritmi di decifrazione che ne consentano e proteggano la *privacy*. In effetti, non è necessario sapere nulla del singolo paziente, ma è necessario sapere *trend* di cura, aderenza alle terapie, *benchmarking* tra ospedali e regioni e un'infinità di dati che solo ora si affacciano all'analisi e alla curiosità dei ricercatori.

In merito alla gestione delle risorse sanitarie, sono comunque in atto esperienze regionali che già da tempo hanno avviato un processo di informatizzazione di più ampio respiro attraverso lo sfruttamento dei *big data*: si tratta di realtà che ora cominciano a raccogliere i frutti di quanto seminato nel recente passato. È il caso del progetto *Acg* (*Adjusted Clinical Groups*) della Regione Veneto, illustrato durante la scorsa edizione di *Health IT 2015*, evento organizzato da IIR - Istituto Internazionale di Ricerca.

Obiettivo dell'iniziativa è stato quello di fornire una mappatura dei variegati bisogni di salute espressi dal territorio, per capire quali siano, nelle varie zone della regione, le tipologie di paziente prevalenti, le malattie più frequenti e, di conseguenza, la quantità e la qualità dei farmaci più impiegati con le relative spese sostenute.

Da questa ricerca sono emerse informazioni utili per la quantificazione dei bisogni sanitari della popolazione, che hanno permesso di migliorare percorsi di cura capaci di generare modelli predittivi epidemiologici in grado di aiutare la Regione Veneto nella distribuzione delle risorse economiche sul territorio.

In sostanza, grazie alla soluzione *Adjusted Clinical Groups* sviluppata dalla *Johns Hopkins University*, la Regione Veneto, sfruttando i *big data*, è in grado, oggi, di classificare la popolazione in gruppi omogenei per tipologia di malattie e consumo di risorse assistenziali, con grandi vantaggi per i cittadini che possono essere assistiti e curati meglio e per chiunque debba prendere decisioni nel contesto sanitario.

La realtà veneta, come funziona

Prendiamo in considerazione la sperimentazione *Acg* veneta come esempio di utilizzo dei *big data* e per dare conto di come, effettivamente, l'utilizzo dei *big data* sia una nuova frontiera. Seguirà anche un accenno alla realtà lombarda che, partendo dagli stessi *big data*, offre una visione lievemente diversa.

Nei primi 3 anni di sperimentazione, il sistema *Acg*, su un campione di due milioni di assistiti, distribuiti in sei ASL della Regione Veneto, ha dapprima testato soprattutto la capacità di misurazione della salute nel territorio, per poi concentrarsi

sulla potenzialità nel gestire la salute nel territorio; obiettivo principale del secondo anno di attività è stato quello non solo di mappare e caratterizzare popolazioni che avessero caratteristiche sociodemografiche più eterogenee rispetto alle popolazioni molto urbane della ULSS 16 di Padova e della ULSS 20 di Verona, analizzate nel corso del primo anno di attività, ma anche di sperimentare il sistema *Acg* come strumento di gestione dell'assistenza sanitaria, soprattutto primaria, nel Veneto.

L'*Acg* è in grado di elaborare i dati a disposizione della Regione, in gran parte provenienti dai flussi informativi delle stesse Usl (ricordiamo, tra le fonti i ricoveri, la specialistica ambulatoriale, la farmaceutica, il Pronto Soccorso e l'assistenza domiciliare), di raggruppare le diagnosi di malattia e l'uso di farmaci in insiemi clinicamente coerenti e, in conclusione, di raggruppare la popolazione in 93 gruppi omogenei per profilo di salute e consumo atteso di risorse assistenziali; questi gruppi vengono, quindi, suddivisi in 6 categorie secondo un criterio strettamente economico e vengono aggregati a quelli a cui sono associate cure similmente dispendiose (Fig. 1).

È interessante notare che il peso economico dei gruppi è estremamente variabile: se si considera che il 20% della popolazione, appartenente alle 3 categorie più critiche, assorbe il 75% delle risorse a disposizione, il dato rende chiaro quanto, per una ASL, possa fare la differenza dover far fronte alle esigenze di persone appartenenti a categorie differenti e quanto, dunque, sia prezioso, per chiunque sia coinvolto nella gestione o nel controllo di tali Aziende sanitarie, avere uno strumento in grado di fornire questo tipo di informazione.

Il sistema *Acg* sfruttando il meccanismo dei raggruppamenti per profilo di salute, è, dunque, in grado di censire la distribuzione delle patologie nel territorio. Utilità primaria di tale funzione è quella di riconoscere la maggiore o minore concentrazione di malattie nel territorio, in modo tale che ogni gruppo di assistiti, misurato nella sua esperienza totale di malattie, singole o multiple, attraverso dati non solo ospedalieri, possa essere assistito meglio. Non solo diventa possibile conoscere i bisogni di salute dei cittadini, ma anche calcolare i rischi legati all'insorgere delle patologie, di modo che, sul piano dell'organizzazione delle strutture, il sistema sanitario potrà, infatti, muoversi in modo preventivo, gestendo al meglio le peculiarità del territorio di competenza.

Il progetto *Acg*, inoltre, una volta identificati i gruppi di soggetti con bisogni assistenziali più complessi, offre l'opportunità di inserire queste persone in programmi professionali di "*case management*" (progettazione e coordinamento dei servizi di assistenza sanitaria adeguati per raggiungere l'obiettivo della riabilitazione medica), che propongono modelli assistenziali innovativi per il territorio e che vedono protagonisti, lavorando in modo sinergico, il medico di medicina generale, il personale infermieristico del territorio che fa riferimento all'ambulatorio del medico e la famiglia dell'assistito.

Infine va nuovamente rimarcato come l'utilità dello stru-

Popolazione e costi

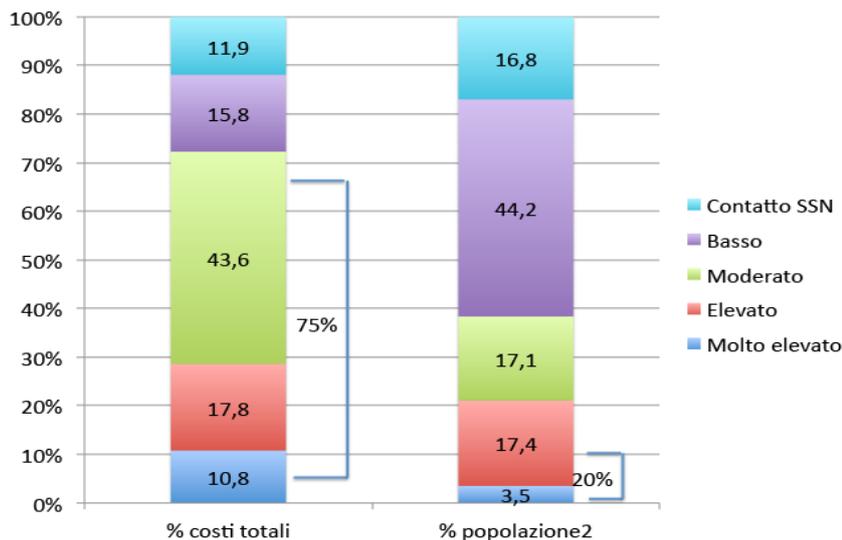


Fig. 1 - Il 20% della popolazione, appartenente alle 3 categorie più critiche, assorbe il 75% delle risorse a disposizione.

mento sia enorme, oltre che sul piano medico, anche e soprattutto su quello della gestione delle risorse, per chi deve prendere decisioni nel contesto della sanità pubblica. Va, inoltre, rammentato che l'*Acg* potrà essere utile per attivare anche meccanismi di controllo sui medici di base; diventa, infatti, possibile verificare con buona approssimazione se il medico che "produce" spese particolarmente alte, ovvero che effettua prescrizioni a cui corrispondono costi superiori alla media, sia effettivamente "giustificato". Basterebbe verificare, per esempio, che i pazienti che ha in carico appartengano effettivamente a gruppi clinici dal peso economico maggiore.

Alcuni in particolare sono gli elementi da mettere in evidenza:

- 1) diversamente da altri strumenti che basano il raggruppamento su episodi di cura (servizi di assistenza sanitaria legati a un certo periodo di tempo, di solito quello della degenza in ospedale e legati a un singolo problema di salute), il sistema utilizzato considera, grazie alla varietà dei dati analizzati, il complesso delle patologie che la persona affronta nel suo percorso, dentro al contesto ospedaliero e fuori, cercando di evitare che l'ospedale divenga l'epicentro del sistema, dal momento che la popolazione sperimenta ormai la maggior parte dei propri percorsi di cura al di fuori di questa struttura. Diventa, infatti, importante incentivare la diversificazione dei luoghi in cui il paziente viene assistito, agevolando anche realtà extra-ospedaliere territoriali;
- 2) In Sanità, la nozione di malattia unica non esiste più perché la norma è diventata la multimorbilità. La popolazione anziana, che poi è il destinatario principale degli interventi di assistenza sanitaria e socio-sanitaria,

non è una popolazione affetta da patologie singole. Nel complesso, come rivelano i dati solo il 50% degli uomini e il 35% delle donne hanno zero o una sola patologia, mentre tutto il resto della popolazione ne ha due o più. Questo, per esempio, potrebbe mettere in crisi alcuni modelli di Hub e Spoke, che diversificano e sistemano in periferia branche e specialità nel momento in cui invece, come dimostrano questi dati, il paziente è comorbido e ha bisogno di specialisti diversi. In altre parole, sistemare l'ortopedia o la chirurgia in periferia per interventi di bassa complessità potrebbe rivelarsi dannoso in pazienti anziani, cardiopatici diabetici e nefropatici, anche se la procedura di intervento dovesse essere di bassa o modesta complessità;

- 3) rispetto al tema dei dati, il sistema *Acg* attinge, come accennato, a svariate fonti disponibili. Una delle sfide, quindi, è quella di riuscire a creare quanti più dati possibili in grado di essere elaborati dal sistema attraverso i quali si potrà rendere la mappatura territoriale dei bisogni di salute sempre più corrispondente alle effettive esigenze palesate dalla popolazione, cercando di ottenere il raggiungimento del maggior grado di benessere bio-psico-sociale.

Il progetto lombardo di *data-warehousing*

In Lombardia è stato avviato un progetto di raccolta di dati sanitari grazie all'utilizzo di *Denali*, un *data-warehouse* sviluppato da Regione Lombardia e Università di Milano-Bicocca, per il monitoraggio e il miglioramento della qualità e dell'appropriatezza delle cure sanitarie fornite ai cittadini lombardi. *Denali* rappresenta uno dei più grandi progetti di *data-wa-*

rehouse al mondo. Le informazioni raccolte in modo anonimo riguardano oltre 9 milioni di residenti in Lombardia tra il 2000 e il 2010. Permette, quindi, di ricostruire il percorso terapeutico attuato da ogni paziente e consente di analizzare tutte le prestazioni sanitarie erogate a livello di macro-aggregati per area di interesse (come, per esempio, patologie, consumo di farmaci, procedure, ospedalizzazioni, ecc.), fornendo importanti valutazioni epidemiologiche per lo sviluppo e il miglioramento della rete assistenziale lombarda. Per le diverse aree terapeutiche prese in esame dal progetto, è stato evidenziato che, per esempio, degli oltre 18.000 pazienti lombardi affetti da diabete mellito e da una grave complicanza come l'arteriopatia periferica, quelli trattati per tale patologia sono meno della metà. Si rileva, quindi, una tardiva presa in carico dei pazienti che, nel 20% dei casi, subiscono, come prima procedura, un'amputazione.

In particolare, l'analisi dei percorsi terapeutici mette in luce che i pazienti sottoposti a procedure di rivascolarizzazione, oltre a mostrare curve di sopravvivenza migliori, hanno tassi di ospedalizzazione inferiori rispetto a quelli che subiscono amputazioni maggiori, con, anche, una degenza complessiva inferiore. Infine, osservando il fenomeno da un punto di vista dei costi, si stima che il costo medio nel periodo di osservazione è di circa 35.000 euro per paziente, mentre l'onere del non trattamento per la popolazione affetta da arteriopatia periferica è di circa 29.000 euro a paziente.

Il Piano nazionale esiti di Agenas

Sviluppato dal governo italiano e dal ministero della salute, le misure del PNE (4) sono strumenti di valutazione a supporto di programmi di *auditing* clinico e organizzativo finalizzati al miglioramento dell'efficacia e dell'equità nel SSN. Nel sito si legge che "PNE non produce classifiche, graduatorie e giudizi" anche se spesso i giornali hanno stilato statistiche e classifiche sugli ospedali peggiori e migliori in Italia. Il cuore del PNE è la SDO e tutte le informazioni deducibili da essa come mortalità, indice di reospedalizzazione, gravità codificata dal sistema ICD9 e così via. La più grande limitazione è l'esclusione degli altri *database* amministrativi che rendono poco profonda l'analisi.

A titolo di esempio, riportiamo la cartina con l'analisi del peggioramento dell'IRC dopo due anni dal primo ricovero. Il sistema è ancora grezzo e non ha le potenzialità dei sistemi veneto e lombardo, ma la strada è ormai tracciata e non ci sono dubbi sul fatto che, nel tempo, avremo sempre più *database* che afferiranno al PNE. Ricordiamo, infatti, che il Ministero ha affidato ad Agenas, tramite, appunto, il PNE, la sorveglianza del SSN.

Altro utilizzo dei *big data* è il *report* OSMED sull'uso dei farmaci (5). Il sistema della farmaceutica è basato sull'utilizzo dei *database* del consumo farmaceutico territoriale, che viene convogliato in un unico *database* e, poi, elaborato. Pur-

troppo, per ora, ci si limita ad analisi farmaco economiche e saltuariamente anche ad analisi di patologia e aderenza, ma è sicuro che, da questo, si potrebbero avere risposte molto utili per la clinica, per dettare risposte e strategie di assistenza.

Conclusioni

L'avvento dell'utilizzo dei *big data* in sanità si pone come un'alternativa, non sappiamo ancora quanto valida, all'epidemiologia. Ma sicuramente le applicazioni sono enormi: incrociare dati per avere informazioni complesse permette di sapere e valutare le interconnessioni tra aspetti diversi spesso studiati separatamente (6). Per esempio, si potrà sapere se il taglio alla farmaceutica provoca spese a carico dei ricoveri o se la riduzione della degenza media provoca un aumento della spesa farmaceutica o della mortalità. Allo stesso modo, si potrà valutare l'impatto che le liste d'attesa hanno sulla salute pubblica in termini di complicanze e spesa, così come quali possano essere gli esiti dei ricoveri effettuati da ospedali piccoli o dai reparti specialistici. Alcune esperienze di utilizzo a fini speciali sono già in fase avanzata, come l'appropriatezza nel diabete della terapia farmaceutica. Sono milioni e milioni i dati da incrociare e siamo solo all'inizio di un'avventura che richiede competenze diverse. Non è certo facile incrociare i dati ed è necessario l'apporto di informatici, epidemiologi, economisti e clinici che possano supervisionare i vari e innumerevoli aspetti. Il desiderio e la possibilità di incrociare i dati potrebbero portare a correlazioni spurie e a interpretazioni errate, ma le analisi si affineranno con il tempo e si potranno stabilire vere e proprie linee di ricerca. Anche in termini occupazionali ci potrebbero essere sbocchi importanti; non tutti hanno le competenze necessarie per utilizzare tali masse di dati e, forse, creare corsi di laurea proprio per l'utilizzo dei dati in sanità non sarebbe certo disdicevole. Un altro aspetto da non sottovalutare è la potenza di tali strumenti soprattutto nell'essere "*super partes*" e non influenzabili dal politico di turno, dal direttore generale o dall'assessore. In altre parole, potrà sempre essere possibile cambiare l'interpretazione da attribuire alle analisi, che, però, proprio per essere disponibili a vari livelli e zone, non potranno certo essere "inquinata" da considerazioni politiche; inoltre, questa interpretazione permetterà analisi di *benchmarking* potenti e asettiche. I dati saranno comunque sempre disponibili e utilizzabili.

Il sistema di *benchmarking*, che permette comparazioni multidimensionali, consentirà di avere le differenze tra ospedali, regioni macroaree e così via. Avremo dati che, anche se non affidabili e non attendibili al 100%, almeno secondo i dettami della letteratura medica, sicuramente lo saranno molto più di piccole serie di dati licenziate dalle varie direzioni aziendali spesso a loro vantaggio. I sistemi di analisi così ottenuti non saranno influenzabili e sbugiarderanno rapporti pieni di parole e pochi fatti. E, soprattutto, si potranno avere dati sui risultati ottenuti e sull'impiego delle risorse.

Potremo sapere gli effetti delle scelte e i libroni aziendali con tanti numeri di cosa si è fatto saranno superati da sistemi che permetteranno di dire non solo cosa e quanto abbiamo fatto (dati banali che ormai nessuno più accetta nell'epoca del *Less is More*)(7), ma come e con quali risultati confrontati con realtà simili in termini di risorse o di numerosità. Potremo tracciare, per esempio, 40.000 dializzati e sapere certamente quante dialisi hanno fatto ma anche quanti ricoveri, quale spesa farmaceutica sia stata sostenuta per ognuno di essi, quanti esami strumentali e, fondamentalmente, l'*outcome*.

In conclusione, è plausibile sostenere che, in un contesto come quello della sanità pubblica, se effettivamente ne vengono utilizzate le reali potenzialità, l'analisi dei *big data* non solo potrà produrre informazioni molto utili in campo clinico ed epidemiologico, ma sarà fondamentale per preservare quel modello di *welfare* assistenziale che caratterizza da sempre la sanità italiana e che, negli ultimi anni, è stato mantenuto con difficoltà dalle Regioni a causa dei tagli e della *spending review* imposti dai governi che si sono succeduti. Sarà possibile tracciare le sacche di inefficienza e spreco. Tutto questo sarà possibile a patto che le nostre amministrazioni riescano (o vogliano), facendo tesoro di esperienze pionieristiche, ma sempre più frequenti, a cogliere le indicazioni estremamente attendibili che emergeranno dalle correlazioni prodotte da questo tipo di analisi.

Alla fine di questo articolo non possiamo esimerci dal segnalare che ognuno di noi è tracciato e, se è lievemente più difficile evitarlo dal telefonino, è sicuramente più facile provare a evitarlo dal *computer* di casa, utilizzando le finestre di navigazione anonima che molti *browser* mettono a disposizione (come *Firefox* e altri) o utilizzando motori di ricerca anonimi. Non servirà, certo, a ingannare sistemi complessi,

ma servirà a non farci inondare di offerte per andare a soggiornare o a mangiare a Pescasseroli semplicemente perché abbiamo solo cliccato su un sito contenente come *metadato* Pescasseroli.

Disclosures

Financial support: No financial support was received for this submission.

Conflict of interest: The authors have no conflict of interest.

Bibliografia

1. Mayer-Schonberger V, Cukier K. *Big data*. Milano, Garzanti Libri S.r.l. 2013.
2. Morozov E. *Internet non salverà il mondo*. Milano, Arnoldo Mondadori Editore 2014.
3. Degli Esposti L, Sturani A, Quintaliani G, Buda S, Degli Esposti E. Administrative databases of the Local Health Unit: possible use for clinical governance of chronic kidney disease. *G Ital Nefrol* 2014; 31 (3); gin/31.3.14.
4. PNE <http://95.110.213.190/PNEed14/>.
5. Degli Esposti L. The Nationwide Osmed Health-Db Database. A Tool To Support Healthcare Decision-Making And Real-World Evidence Generation. *Value Health* 2015; 18 (7): A693. doi: 10.1016/j.jval.2015.09.2577. Epub 2015 Oct 20.
6. Degli Esposti L, Saragoni S, Buda S, et al. Awareness of albuminuria in an Italian population-based cohort of patients treated with hypoglycemic drugs. *J Nephrol* 2012; 25 (3): 325-31.
7. PNE <http://95.110.213.190/PNEed14/>.