

La grandezza risiede nella partecipazione

La condivisione dei dati individuali sanitari può migliorare la salute della società e anche della ricerca

Navigando sul web assistiamo sempre più spesso alla generazione di contenuti che sono perfettamente cuciti sulle nostre preferenze. La ricerca di Google traccia i nostri percorsi per comprendere quali tipologie di pagine e documenti apriamo più di frequente. Ogni sito o piattaforma ci propone la pubblicità degli oggetti che preferiamo. I social network che frequentiamo tracciano intorno a noi una rete tridimensionale, che evidenzia le nostre relazioni e ci inserisce all'interno di categorie rilevabili all'interno di una popolazione¹. La capacità predittiva o analitica di internet scaturisce dall'enorme ammontare di dati che semina il web, quando visitiamo i siti o esprimiamo le nostre preferenze con un click del mouse o lasciamo volontariamente i nostri dati ai gestori di piattaforme. Ormai il web permea ogni ambito della nostra esistenza e quindi sono molteplici le occasioni di produzione ma soprattutto di raccolta di dati.

Anche la medicina non è immune alla raccolta di informazioni sui pazienti. Anzi la clinica sembra avere già a disposizione possibili sistemi per la collezione di dati provenienti dai pazienti, tra l'altro senza dover spendere cifre esose per ottenerli. Infatti sono i pazienti stessi a generare dati nella loro quotidianità, ad esempio nel momento in cui usufruiscono delle strutture sanitarie o quando sfruttano strumenti elettronici per monitorare i propri parametri o comportamenti, al fine di migliorare la propria condizione di salute. L'introduzione della cultura del dato nella ricerca biomedica o nella pratica clinica produrrà conoscenza sia per i medici sia per i pazienti, mentre a livello della ricerca potrebbe dare impulso a commistioni di idee e tecniche.

Tuttavia l'utilizzo di questi dati, finora ancora poco sfruttati, non deve sfuggire a un'attenta riflessione. Secondo la Shared nationwide interoperability roadmap, la condivisione dei dati dovrebbe aiutare a mettere il paziente al centro di tutto il processo di cura². Il confronto con l'evoluzione che il fenomeno sta avendo sul web ci mette in guardia: il rischio è di creare una grande mole di informazioni in continuo movimento, senza che le persone siano poste al centro di questo

flusso. L'uso e l'elaborazione dei dati raccolti dal web spesso non avvengono alla luce del sole e l'informazione generata non è fruibile dall'utente iniziale, colui che di fatto ha generato questa immensa ricchezza.

Ecco perché è così urgente una riflessione sui big data in ambito sanitario. Se rispetto ad altri ambiti della scienza la medicina è ancora un passo indietro nella condivisione dei dati (si pensi ad esempio al progetto genoma o ai dati ottenuti dal telescopio spaziale Hubble), occorre cogliere questo ritardo come l'occasione per superare le barriere che un'informazione così personale, come quella sulla salute, introduce e come l'opportunità per costruire strutture e infrastrutture capaci di garantire accesso e coinvolgimento tanto ai pazienti quanto al personale sanitario.

“ Occorre sviluppare una cultura per cui la condivisione di un dato personale rappresenti un beneficio per l'intera comunità. ”
— Harlan Krumholz

L'accesso ai dati implica partecipazione

Viviamo in un'epoca in cui sono attivi movimenti quali il crowdsourcing o la sorveglianza partecipativa, e dove il fatto di donare dati viene percepito come un impegno individuale, addirittura attraverso la quantificazione di parametri fisiologici tramite dispositivi elettronici. Questo fermento andrebbe sfruttato per contribuire a una maggiore conoscenza delle malattie, a una diagnosi sempre più puntuale e a cure e trattamenti più personalizzati. La ricerca, l'analisi e la sistematizzazione delle informazioni generate dalla popolazione sono una ricchezza per la società tutta, ma sono anche facile preda di mire commerciali. E le persone se ne stanno accorgendo. Lo dimostra l'iniziativa Care data³ condotta nel Regno Unito, che ha riscosso scarso successo a causa del fatto che le persone non hanno condiviso l'idea che le proprie informazioni sanitarie venissero cedute alla ricerca e, soprattutto, ad aziende concentrate sul profitto.

Per non perdere l'accesso ai dati relativi alla salute, c'è un conflitto che va risolto tra il diritto di preservare la privacy e l'opportu-

nità di una sanità guidata da nuove evidenze scientifiche che possono scaturire dall'analisi di una grande quantità di dati. Se da un lato nessuno sembra porsi il problema di cedere informazioni personali a piattaforme o dispositivi elettronici, la scarsa comunicazione da parte delle istituzioni, le agende poco chiare e l'insufficiente trasparenza sul controllo e la proprietà dei dati relativi alla salute, creano imbarazzo nel momento in cui viene richiesta la loro condivisione a soggetti terzi.

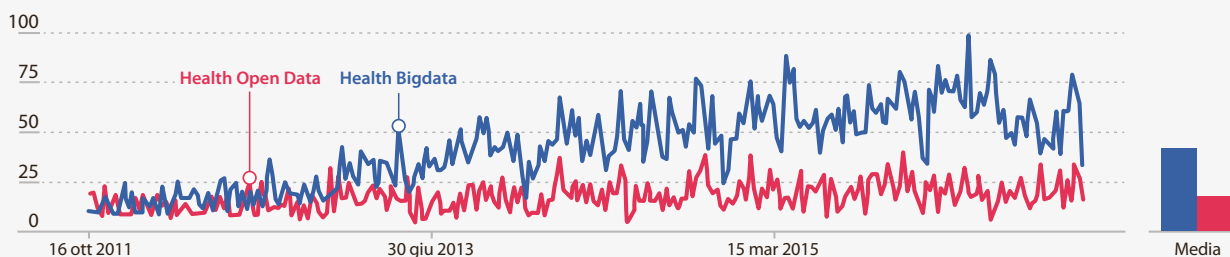
La discrepanza cui assistiamo non è questione di ingenuità dei cittadini ma di impostazione della richiesta. Occorre pensare attentamente alla formulazione del consenso, perché il cittadino non abbia l'impressione di perdere il controllo dei dati personali, una volta che si è rivolto a un ente per un proble-

ma di salute. La mancanza di trasparenza in fase di condivisione o elaborazione del dato potrebbe destare sfiducia e sospetti nei confronti delle istituzioni. Diverso è il discorso per le piattaforme e i dispositivi elettronici, dove il cittadino acconsente a "donare" i propri dati, nel momento in cui crea un proprio account e sottoscrive termini e condizioni di utilizzo⁴.

Per superare la diffidenza, occorre incoraggiare il coinvolgimento dei cittadini nel processo di produzione, di gestione e di fruizione dei dati. Dialogo e trasparenza sono gli elementi fondanti per recuperare la fiducia, perché l'assenza di dibattito elimina la possibilità di costruire una consapevolezza. Occorre chiarire e rafforzare il concetto dei diritti legati alle proprie informazioni personali, perché sia piena la comprensione dei rischi e dei benefici legati alla condivisione dei dati. Sono inoltre incoraggianti le storie di successo, che aiutano a mettere in evidenza l'utilità della condivisione dei dati. Oltre al coinvolgimento, occorre assicurare sistemi capaci di rafforzare la trasparenza nel processo di amministrazione del dato e di garan- a p.14 →

La ricerca dei trends di Google, focalizzata sugli ultimi cinque anni, evidenzia che cresce l'interesse nei confronti dei big data relativi alla salute, mentre quello per gli open data mantiene un andamento costante. Il tema dell'accessibilità ai dati dovrebbe essere maggiormente dibattuto.

La passione per i big data



da p.13 → tire la gestione dei dati in modo anonimo. L'introduzione di consensi informati che mettano l'utente nelle condizioni di scegliere i destinatari dei propri dati sarebbe di grande vantaggio. Così come l'introduzione di meccanismi di notifica in grado di informare l'utente di possibili violazioni.

Altro elemento rassicurante potrebbe essere un sistema legislativo⁵ più omogeneo, dato che le diverse scelte adottate da Stati Uniti e Europa sono disorientanti.

"Occorre sviluppare una cultura per cui la condivisione di un dato personale rappresenti un beneficio per l'intera comunità", ha affermato Harlan Krumholz in un'intervista al *New England Journal of Medicine*⁶. "Nel processo di condivisione ogni persona deve essere messa in grado di vedere il dato, comprenderlo ed eventualmente usarlo."

La partecipazione declinata nel contesto della ricerca clinica

È sulla base di tale prospettiva che nel 2011 Krumholz ha avviato il progetto Yale open data access (Yoda) perché nel caso della ricerca biomedica non mancano solo gli standard ma spesso anche le piattaforme capaci di garantire al contempo privacy, accessibilità, condivisione e riconoscimento degli autori della ricerca. Lo Yoda nasce proprio per rispondere a questi bisogni, presentandosi come un partner indipendente che fornisce supporto e assicura imparzialità e trasparenza a chiunque voglia condividere i dati di trial clinici. I principi guida della piattaforma sono il rispetto della privacy dei partecipanti coinvolti, mediante la sottoscrizione di un consenso informato. Chiara è la dichiarazione dei fini per cui i dati vengono raccolti: i progetti ammessi alla piattaforma Yoda per la condivisione dei dati hanno solo fini scientifici, volti a migliorare la salute pubblica e rifiutano qualsiasi utilizzo commerciale. Il principio cardine è la trasparenza che impone la tracciabilità delle decisioni e la condivisione dei risultati. La trasparenza, infatti, è garantita proprio dall'accesso all'elaborazione delle informazioni e dalla possibilità di condivisione. "Bisogna superare la cultura per cui solo il produttore del dato è in grado di capire come usarlo e interpretarlo", ha continuato Krumholz. Per questo motivo è necessario accompagnare il dato con metadati che spieghino come un risultato è stato ottenuto e come può essere interpretato.

Yoda è dunque il tentativo di proporre uno standard accessibile a vari livelli, capace di assicurare l'interoperabilità nella gestione delle banche dati. Ciò si traduce nella realizzazione all'interno del progetto di una piattaforma utile all'elaborazione dei dati e predisposta per una sicura condivisione dei risultati. "Dobbiamo offrire una piattaforma affidabile ed essere responsabili nella gestione dei dati, per dimostrare a tutti che questa azione è volta a migliorare la società", ha spiegato Krumholz. Dunque la partecipazione è significativa anche a un altro livello, quello della gestione del dato e del coinvolgimento del personale sanitario e dei medici.

Per procedere a più larghi passi la ricerca ha bisogno dei dati. Lo dimostrano le richieste di dati avanzate nei confronti dell'industria farmaceutica, che ha risposto in alcuni casi predisponendo programmi di condivisione delle informazioni cliniche e sperimentali raccolte. Questo tipo di impostazione è rafforzata dall'International committee of medical journal editors (Icmje) che sta tentando

di diffondere una cultura basata sul concetto di open science, chiedendo ai suoi autori di condividere i dati generati dai trial clinici e di considerarlo come un "dovere morale"⁷. Secondo Krumholz, tale richiesta potrebbe inaugurare una tendenza culturale, capace perfino di spingere i pazienti a una maggiore partecipazione agli studi clinici.

Tuttavia, affinché i big data possano apportare benefici reali alla clinica e alla società intera è indispensabile modificare le impostazioni di ricerca. Il metodo scientifico finora applicato in medicina prevede il più delle volte di formulare un'ipotesi che deve essere in seguito verificata. Qui invece si tratterebbe di partire dai dati raccolti dall'osservazione dei pazienti, per poi formulare una spiegazione del fenomeno. Il ragionamento deduttivo verrebbe quindi sostituito da quello induttivo, un'impostazione che spaventa, perché potrebbe introdurre risultati falsi positivi. Per scongiurare tale possibilità occorre un robusto sistema di validazione dei dati, soprattutto quando si vogliono ricreare delle relazioni causali. Appellarsi ai principi della open science e alla condivisione dei risultati diventa quindi cruciale per validare i risultati tramite la replicazione, che non può avvenire



“Bisogna superare la cultura per cui solo il produttore del dato è in grado di capire come usarlo e interpretarlo.”

— Harlan Krumholz

se i dati non sono aperti e se non c'è la possibilità di verificare la riproducibilità di quanto affermato dai risultati ottenuti in differenti condizioni e coinvolgendo diversi gruppi di pazienti⁸. Oggi la maggior parte dei dati prodotti dai trial clinici non passa attraverso la revisione della peer review. Per fare un esempio che descrive il fenomeno, l'88% degli autori di articoli comparsi su giornali finanziati dai National Institutes of Health statunitensi non deposita i propri dati all'interno degli appositi registri. Come conseguenza questa cultura scientifica è inefficiente e irripetibile. Secondo Eric L. Uhlmann e Raphael Silberzahn questo meccanismo – oltre a rallentare il progresso scientifico – sta sprestando gli investimenti pubblici e disonora chi contribuisce alla ricerca⁹.

Senza contare che rilasciare i dati in formato aperto è un beneficio tanto per l'ambito clinico quanto per la ricerca, perché permette di correlare tra loro informazioni differenti in un sistema definito *linked data*, che può includere anche elementi provenienti dal mondo della biologia, dell'economia o dalle caratteristiche demografiche o ambientali: un tale grado di complessità è ciò che contraddistingue i big data. La correlazione semantica di

dati provenienti da diverse fonti consente di creare corrispondenze tra sintomi, malattie, diagnosi, trattamenti e prescrizioni, ampliando la conoscenza e permettendo di produrre nuove evidenze¹⁰.

Il fatto di rendere disponibili i dati è la chiave per trovare relazioni tra fenomeni che altrimenti difficilmente sarebbero identificabili. Occorre dunque permettere lo scambio e in seguito l'armonizzazione di tutti i dati legati alla salute, che risiedono in luoghi diversi e che sono archiviati in differenti formati¹¹. Combinare set di dati che non sono basati su modelli comuni, a volte generano duplicati o informazioni conflittuali e in generale possono introdurre problemi nella pratica.

La possibilità di accedere ai dati non è dunque sufficiente per creare benefici. Bisogna avanzare nella tecnica e nelle capacità analitiche e sviluppare nuove applicazioni. Per maneggiare la varietà e il volume di informazioni che caratterizzano i big data non si può prescindere dalla collaborazione tra diverse figure professionali, provenienti sia dal mondo del machine learning e del data mining sia dall'ambito tipicamente biomedico, che ha gli strumenti per interpretare e valutare i risultati.

Accanto alle competenze servono piattaforme che rendano questi dati facilmente condivisibili e maneggevoli. Le applicazioni per gestire i dati devono essere semplificate affinché diventino accessibili a tutti⁸. L'Hit strategic plan per il 2015–2020 rafforza il concetto dell'accessibilità ponendo l'obiettivo di "migliorare la salute del singolo individuo e delle comunità attraverso l'uso delle informazioni derivate dalla tecnologia, che devono essere accessibili se rilevanti per più persone"¹².

Qual è dunque la possibilità del singolo o delle comunità, anche non scientifica, di interfacciarsi con i big data? In definitiva, osservati dalla prospettiva dei cittadini, essi potrebbero migliorare la comprensione di una malattia, perfezionare le terapie per le patologie croniche oppure offrire tutti gli elementi per curarsi da casa. Allargando lo sguardo al sociale, sarebbero l'occasione per sottolineare le disuguaglianze o le necessità proprie di alcune aree. I cittadini e le comunità locali, se opportunamente istruiti, potrebbero trovare risposte per problemi locali e disegnare possibili soluzioni¹³.

Occorre creare delle metriche riportate pubblicamente per assicurare un facile accesso ai dati per i pazienti che si rivolgono agli ospedali o ai sistemi sanitari. Oggi, a volte, non sono nemmeno accessibili le note inserite nella cartella clinica, secondo quanto sostiene il movimento OpenNotes. C'è ancora molto da fare nella realizzazione di piattaforme che promuovano la ricerca, l'accesso, la possibilità di operare e infine il riuso dei dati¹³.

È un'opportunità unica per le persone senza una preparazione medica il fatto di avere dati di confronto con altre centinaia di malati che hanno caratteristiche cliniche simili e che hanno provato determinati trattamenti. Con queste informazioni è possibile porre basi solide per una discussione basata su evidenze scientifiche con i clinici¹⁴. Se le persone potessero esprimere la propria opinione su come i loro dati vengono utilizzati, per garantire migliore qualità, informazioni più aggiornate e più complete, probabilmente si arriverebbe veramente ad aiutare tutti quegli individui che si trovano ad affrontare problemi simili.

[Pagine a cura di **Giulia Annovi**]

I cinque obiettivi dell'Health It strategic plan 2015-2020 statunitense

RACCOLTA

Primo obiettivo

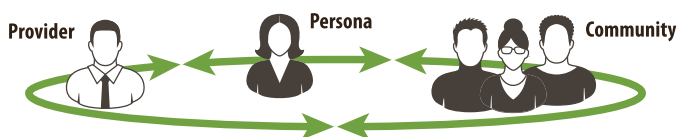
Espandere l'adozione dell'information technology in sanità.



CONDIVISIONE

Secondo obiettivo

Migliorare la sicurezza dei dati sanitari e l'interoperabilità dei sistemi.



UTILIZZO

Terzo obiettivo

Rafforzare il trasferimento all'assistenza sanitaria.



Quarto obiettivo

Migliorare la salute e il benessere delle persone e delle comunità.



Quinto obiettivo

Migliorare l'innovazione e la conoscenza della ricerca scientifica.



- Lazer D, Pentland A, Adamic L, et al. Social science. Computational social science. Science 2009;323:721-3.
- The office of the national coordinator for Health information technology. Connecting Health and Care for the Nation. A Shared Nationwide Interoperability Roadmap. Final version 1.0, 2015.
- Goldacre B. Care.data is in chaos. It breaks my heart. The Guardian 2014; 28 febbraio.
- Kostkova P, Brewer H, de Lusignan S, et al. Who owns the data? Open data for healthcare. Front Public Health 2016; 4:7.
- Lupton D. Health promotion in the digital era: a critical commentary. Health Promot Int 2014; 30:174-83.
- Krumholz HM, Waldstreicher J. The Yale open data access (Yoda) project – A mechanism for data sharing. N Engl J Med 2016;375:403-5.
- Taichman DB, Backus J, Baethge C, et al. Sharing clinical trial data: a proposal from the International committee of medical journal editors. Ann Intern Med 2016; 64: 505-6.
- Krumholz HM. Big data and new knowledge in medicine: the thinking, training, and tools needed for a learning health system. Health Affairs 2014;33:1163-70.
- Silberzahn R, Uhlmann EL. Crowdsourced research: many hands make tight work. Nature 2015;526:189-91.
- European Commission. The socio-economic impact of interoperable electronic health record (Ehr) and ePrescribing systems in Europe and beyond. 2009.
- Murugiah K, Ritchie JD, Desai NR, et al. Availability of clinical trial data from industry-sponsored cardiovascular trials. J Am Heart Assoc 2016;5:e003307.
- The Office of the national coordinator for health information technology. Federal health It strategic plan: 2015-2010. 2014.
- Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The Fair guiding principles for scientific data management and stewardship. Sci Data 2016;3:160018.
- Krumholz HM, Terry SF, Waldstreicher J. Data acquisition, curation, and use for a continuously learning health system. Jama 2016; doi: 10.1001/jama.2016.12537

VEDI ANCHE

Condividere, governare e utilizzare i dati

favorisce un sistema sanitario capace di apprendere costantemente dall'esperienza: ne è convinto Harlan Krumholz che è tornato a esporre il proprio punto di vista in una nota pubblicata nel numero speciale del Jama curato dall'Institute of medicine statunitense¹. Purtroppo, però, questo scenario è ostacolato da diversi fattori: in primo luogo, chi detiene il dato (il più delle volte il ricercatore) non è incentivato a condividerlo, soprattutto perché su di lui (o sulla sua istituzione) ricade la totalità dei costi necessari per l'attività di sharing. Inoltre, l'interoperabilità dei dati è ancora un miraggio. In definitiva, i problemi sono di ordine etico e di natura tecnologica: serve una cultura nuova, in cui la non condivisione delle informazioni sia vista come un oltraggio al progresso. •

- Krumholz HM, Terry SF, Waldstreicher J. Data acquisition, curation, and use for a continuously learning health system. Jama 2016;316:1669-70.

