

Open science: come promuoverla a tutti i costi

Wilhelm EE, Oster E, Shoulson I

Approaches and costs for sharing clinical research data

JAMA 2014; 311 (12): 1201-1202

La produzione, diffusione e condivisione dei dati della ricerca clinica sono gli ingredienti chiave che contribuiscono al progresso scientifico e, di conseguenza, al miglioramento della salute pubblica. In particolare, la condivisione dei dati viene incoraggiata per promuovere la cosiddetta *open science*, ovvero la maggiore trasparenza ed accessibilità della ricerca allo scopo di migliorare lo sviluppo di farmaci e dispositivi medici¹. Nonostante i benefici reali, la condivisione dei dati rappresenta un processo tuttora complesso. I ricercatori, infatti, tendono in genere a custodire gelosamente i dati della ricerca per mantenerne la proprietà intellettuale, ostacolare la concorrenza, proteggere la riservatezza e la privacy, o evitarne l'uso improprio da parte di persone non qualificate. Non ultimo, la condivisione dei dati della ricerca presenta un conto piuttosto salato a chi decide di farsene carico.

I MODELLI

Esistono diversi modelli che consentono di condividere dati; essi variano a seconda dell'accessibilità ai dati, del tipo e dimensione dei dati raccolti, della complessità del database e della predisposizione e capacità dell'interfaccia utente. Nei modelli più semplici i dati dei partecipanti allo studio, una volta resi anonimi, vengono caricati su un sito web al quale i ricercatori possono accedere attraverso username e password. Questo modello, tuttavia, consente un utilizzo limitato del dato da parte dei ricercatori. Modelli più complessi, invece, sono in grado di condividere dati standardizzati, provenienti da più sistemi informativi e accompagnati da una interfaccia *user-friendly*, per renderli maggiormente comprensibili, nonché più facilmente utilizzabili dagli utenti.

Il National Database for Autism Research², la Parkinson's Progression Markers Initiative³ e l'Alzheimer's Disease Neuroimaging Initiative –



ADNI⁴ costituiscono alcuni esempi di modelli complessi, specifici per patologia. Creato nell'ottobre 2004 e attualmente co-finanziato fino al 2017 da governo, industria, sponsor e fondazioni, l'ADNI rappresenta forse la più conosciuta tra queste iniziative e fornisce dati standardizzati per 3 diversi database (clinico, di diagnostica per immagini e di biomarcatori).

L'ADNI permette di condividere in tempi brevi grandi quantità di dati clinici che, una volta acquisiti e inseriti nei database, vengono controllati per garantirne gli standard di qualità e resi accessibili al ricercatore in pochi giorni, attraverso un sito web protetto da password. Dal canto loro i ricercatori che accedono ai dati dell'ADNI non solo si impegnano a mantenere anonimi i partecipanti allo studio, ma anche ad inserire nel database ogni nuova analisi effettuata e a citare l'ADNI in ogni relativa pubblicazione. Ad oggi sono più di 4000 i ricercatori che hanno avuto accesso ai dati dell'ADNI, con il risultato di un totale di 411 pubblicazioni.

I COSTI

Nonostante la sua funzionalità ed utilità nel campo della ricerca clinica, questo modello di condivisione dei dati presenta costi notevoli, che possono essere raggruppati in quattro principali categorie.

- Costi di infrastrutture ed amministrazione: legati al mantenimento degli archivi informativi, alle procedure di sicurezza e controllo della qualità del dato, nonché all'acquisizione del dato stesso, tramite consenso informato del partecipante allo studio clinico;
- Costi per la standardizzazione dei dati: legati al processo di standardizzazione dei dati, per consentirne la condivisione e semplificarne l'utilizzo per le analisi quantitative;
- Costi per le risorse umane: rappresentano la principale voce di spesa, soprattutto se trat-

tasi di modelli di condivisione basati su tecnologie avanzate e sofisticate, la cui gestione e manutenzione richiedono personale qualificato;

- Costi di opportunità: legati al tempo e alle risorse umane utilizzate per la condivisione dei dati, che altrimenti sarebbero state utilizzate per condurre nuovi studi di ricerca.

Sebbene ogni progetto di condivisione di dati clinici presenti caratteristiche ed obiettivi specifici, tutte le condivisioni dei dati comportano dei costi, commisurati alla quantità e qualità dei dati stessi. Come ricordato in precedenza, la modalità più semplice di condivisione, in cui dati non standardizzati sono resi disponibili tramite un sito web, ha sì un costo relativamente basso, ma permette un utilizzo limitato dei dati da parte del ricercatore. I costi della condivisione aumentano in modo direttamente proporzionale alla tempistica, accessibilità e fruibilità del dato. Tali costi, nella maggior parte dei casi sostenuti principalmente dai ricercatori che raccolgono il dato, potrebbero essere recuperati, ad esempio, attraverso il pagamento di un canone di abbonamento per l'accesso ai dati, o dovrebbero perlomeno essere considerati in fase di bilancio dagli sponsor della ricerca, data l'elevata richiesta di dati non soltanto da parte loro, ma anche da parte delle riviste scientifiche e dei ricercatori stessi^{5, 6}.

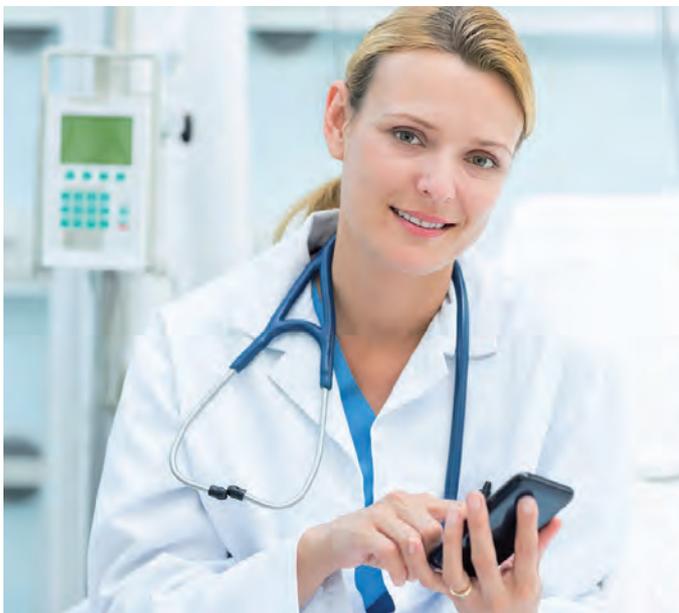
L'assenza di standard di riferimento per una stima attendibile dei costi di condivisione gioca tuttavia un ruolo fondamentale in questo processo; pertanto occorre promuovere una valutazione sistematica di tali costi, che consentirebbe a sponsor, sia pubblici sia privati, di pianificare il budget di spesa in fase preliminare e contribuirebbe indirettamente a utilizzare al meglio il potenziale della condivisione dei dati clinici.

Elia Ferroni

Dipartimento di Epidemiologia, Servizio Sanitario Regionale del Lazio

BIBLIOGRAFIA

1. Ross JS, Krumholz HM: Ushering in a new era of open science through data sharing: the wall must come down. *JAMA* 2013; 309 (13): 1355-1356.
2. National Database for Autism Research (NDAR) data distribution. NDAR website: <http://ndar.nih.gov>. Accessed January 28, 2014.
3. Parkinson's Progression Markers Initiative (PPMI) website: <http://www.ppmi-info.org>. Accessed January 28, 2014.
4. Alzheimer's Disease Neuroimaging Initiative (ADNI) website: <http://www.adni-info.org>. Accessed January 28, 2014.
5. Spertus JA: The double-edged sword of open access to research data. *Circ Cardiovasc Qual Outcomes* 2012; 5 (2): 143-144.
6. Berman F, Cerf V: Science priorities: who will pay for public access to research data? *Science* 2013; 341 (6146): 616-617.



www.recentiproggressi.it

Indicizzata su Medline e Scopus. Ricercabile su PubMed. Accessibile online dagli operatori delle aziende sanitarie e ospedaliere di diverse Regioni italiane. Strumento di consultazione del personale di gran parte degli IRCCS del nostro Paese.

Ma anche utile e indipendente.

Per di più, sul sito della rivista trovi

- videointerviste collegate con i contenuti pubblicati dalla rivista
- infografiche di approfondimento.

Puntuale, credibile, italiana.

