

Valutare l'impatto della ricerca: misurare il guadagno ed evitare lo spreco

Luciana Ballini

Agenzia Sanitaria e Sociale Regionale, Regione Emilia-Romagna; Azienda USL di Reggio Emilia

Riassunto. Il numero di ottobre 2015 di *BioMedCentral Medicine* pubblica un articolo che riporta alcune analisi preliminari del nuovo programma del Regno Unito di valutazione della ricerca, denominato *Research Excellence Framework* (REF) del 2014, primo esercizio di misurazione dell'impatto della ricerca svolta in ambito accademico. Il programma nasce dall'interesse crescente nella possibilità di dimostrare il ritorno di investimento nella ricerca e di ridurre lo spreco derivato dal mancato utilizzo dei risultati della ricerca. In questo contributo una panoramica dei sistemi di valutazione adottati nel Regno Unito dal 2006 in poi, che ne mette in evidenza criticità e punti di forza.

Parole chiave. Ricerca clinica, sprechi, valutazione.

Abstract. The number of *BioMedCentral Medicine* issued in October 2015 published an article encompassing some preliminary analyses of the new UK *Research Excellence Framework* (REF) program developed in 2014 and aimed at assessing research. It is the first exercise conceived to measure the impact of the research carried out in the academic scope. The program is the result of the increasing interest in the possibility of demonstrating a return on investments in the research area and of reducing waste deriving from the non-usage of the research results. This paper is an overview of the assessment systems adopted in the United Kingdom from 2006 forth and it underlines criticalities and strengths.

Key words. Assessment, clinical research, waste.

Il numero di ottobre 2015 di *BioMedCentral Medicine*¹ pubblica un articolo che riporta alcune analisi preliminari del nuovo programma del Regno Unito di valutazione della ricerca, denominato *Research Excellence Framework* (REF) del 2014, primo esercizio di misurazione dell'impatto della ricerca svolta in ambito accademico. Il programma nasce dall'interesse crescente nella possibilità di dimostrare il ritorno di investimento nella ricerca e di ridurre lo spreco derivato dal mancato utilizzo dei risultati della ricerca.

Il dibattito tra ricercatori, finanziatori e utilizzatori finali dei risultati su come definire e rilevare l'impatto della ricerca è presente da tempo, come anche la preoccupazione nei confronti della ricerca inutilmente replicata o con scarsa rilevanza. Le due prospettive sono infatti strettamente legate: per evitare lo spreco occorre avere chiaro quale sia il guadagno.

Il modello più popolare di valutazione di impatto della ricerca è quello lineare, che stabilisce un legame diretto tra ricerca e impatto dei risultati. Il modello è declinato in maniera diversa a seconda dei

punti di vista: i ricercatori tendono a presumere un effetto diretto di uno studio, attraverso la pubblicazione o altri tipi di comunicazione dei risultati; i decisori tendono a presumere che, se adeguatamente commissionata da loro stessi, la ricerca è capace di proporre soluzioni ai loro problemi. Questa visione, qui molto semplificata, è stata messa in discussione dalla letteratura sulla trasmissione efficace della conoscenza (*knowledge translation*) che evidenzia:

1. come la semplice diffusione sia un mezzo meno adeguato dei canali informali, attraverso i quali la conoscenza si sviluppa, viene condivisa e utilizzata all'interno delle comunità di pratica;
2. quanto l'allineamento tra i quesiti dei decisori e i modi/tempi con cui i ricercatori li affrontano richieda sistemi di relazione capaci di colmare il divario concettuale e operativo tra i due settori.

Nel Regno Unito la valutazione della ricerca accademica (*Higher education*) è iniziata negli anni '80 e da allora si sono susseguite diverse modalità di valutazione (e quindi di rendicontazione da parte di istituti di ricerca e ricercatori) sempre più articolate e sempre più gravose dal punto di vista amministrativo.

Nel 2006 il governo inglese decise di sostituire il sistema allora in uso con uno basato su misure bibliometriche, nella convinzione che il cambiamento

¹Greenhalgh T, Fahy N (2015): Research impact in the community-based health sciences: an analysis of 162 case studies from the 2014 UK Research Excellence Framework, *BMC Medicine*, 13: 232.

sarebbe stato ben accetto al settore di riferimento, una volta dimostrato che la bibliometria fosse in grado di misurare la qualità della ricerca e che il parere degli esperti sarebbe stato utilizzato ogni qualvolta queste misure di qualità non fossero state disponibili.

Nel 2007 venne richiesto all'*Higher Education Funding Council for England* (HEFCE), che rappresenta i quattro enti di finanziamento della ricerca accademica del Regno Unito (Inghilterra, Scozia, Galles e Irlanda del Nord), di sviluppare il nuovo modello. Consapevole del fatto che, oltre all'obiettivo principale di informare la struttura di finanziamento e il sistema di assegnazione dei fondi per la ricerca, il modello proposto sarebbe stato usato anche per definire standard di riferimento e per effettuare comparazioni, con ricadute sulla reputazione degli istituti, sulla gestione e programmazione dell'attività di ricerca e sulla carriera accademica dei singoli ricercatori, l'HEFCE diede avvio ad un lungo e complesso processo di consultazione. Il processo si è concluso nel 2010 con la condivisione dei parametri principali, ma anche con una sostanziale modifica – quasi uno stravolgimento – degli indirizzi sottostanti alla proposta politica.

Nel programma di lavoro dell'HEFCE vennero indicati gli obiettivi del nuovo sistema per l'eccellenza della ricerca (REF 2014): definire indicatori per l'eccellenza robusti e applicabili a tutte le discipline, utilizzabili per un confronto con gli standard internazionali, per una assegnazione dei fondi di ricerca che privilegi l'eccellenza, per ridurre il carico amministrativo, per disincentivare comportamenti scorretti, per promuovere parità e diversità, per assicurare il sostegno continuo ad una ricerca leader nel mondo. Un'ulteriore considerazione – esplicitata fin dall'inizio – è stata che il sistema avrebbe dovuto saper riconoscere la qualità e l'eccellenza sia nella ricerca libera (*curiosity-driven*) che in quella programmata in risposta alle richieste degli utilizzatori della ricerca (*agenda-driven*).

Uno dei primi aspetti affrontati – attraverso gruppi di lavoro, consultazioni allargate e studi pilota – ha riguardato l'utilizzo di indicatori bibliometrici. Il risultato dell'analisi – dopo aver evidenziato i numerosi problemi e limiti di questi strumenti – ha portato ad esprimere un consenso sul fatto che gli indicatori di citazioni bibliografiche non sono sufficientemente robusti e non possono sostituire la revisione da parte degli esperti. Inoltre è stato chiaramente espresso, e documentato dallo studio pilota, che gli indicatori bibliometrici comportano una significativa discriminazione di genere e possono incoraggiare strategie e comportamenti votati a migliorare la prestazione bibliometrica piuttosto che la qualità della ricerca. I lavori sono proseguiti con la ricerca di una

definizione condivisa di *impatto* della ricerca, che tenesse conto del suo valore in termini di utilizzo dei risultati mantenendo un'attenzione alla qualità. È stato pertanto proposto che il processo si avvallesse del giudizio di esperti con il coinvolgimento degli utilizzatori finali, al fine di evitare che parametri basati esclusivamente su criteri di qualità accademica finissero per scoraggiare la ricerca intesa a beneficio degli utilizzatori. A seguito dei risultati della prima consultazione, il Segretario di Stato nel 2009 ha richiesto all'HEFCE di sviluppare un sistema che tenesse conto dell'impatto della ricerca sull'economia e sulla società, rendendo chiaro che il sistema non sarebbe stato sostituito da misure bibliometriche. Con il mandato di focalizzarsi sull'impatto della ricerca – e non solo sulla qualità – è iniziato il secondo giro di consultazioni, che ha compreso la committenza alla RAND Europe di uno studio sui diversi approcci esistenti per valutare l'impatto della ricerca.

Il rapporto della RAND (*Capturing research impact. A review of international practice*) fornisce un dettagliato resoconto dell'analisi di quattro casi studio condotti in contesti con una consolidata politica di valutazione della ricerca e delle attività accademiche più in generale (Australia, Regno Unito, Stati Uniti e Paesi Bassi). I quattro sistemi selezionati sono mirati a valutare in maniera aggregata l'impatto di interi programmi di finanziamento della ricerca e la loro descrizione è accompagnata da un ricco dibattito sulle diverse definizioni e modalità di misurazione di impatto.

Un efficace modello di valutazione dell'eccellenza della ricerca doveva rispondere, secondo l'HEFCE, ai seguenti cinque requisiti:

1. essere ritenuto credibile e accettabile dal settore dell'accademia e degli utilizzatori della ricerca;
2. comprendere una vasta gamma di significati di impatto (economico, sociale, sulle politiche e i servizi pubblici, sul welfare, sulla cultura e sulla qualità della vita);
3. essere applicabile a tutte le discipline;
4. essere pratico e non troppo gravoso;
5. evitare percezioni errate e comportamenti indesiderati.

Il modello che ha soddisfatto questi requisiti è quello sviluppato, ma ancora non pienamente adottato, dal governo australiano: esso propone un equilibrio tra impatto e qualità affinché né la ricerca di base né quella applicata siano penalizzate. Si definisce iniquo per definizione – in quanto finalizzato a premiare l'eccellenza – ma equanime nella sua applicazione ad ogni ambito e disciplina di ricerca. La sua unità di analisi non sono i singoli progetti di ricerca bensì i programmi di finanziamento o pro-

grammi di ricerca sviluppati da reti di ricercatori, comunità di ricerca o dipartimenti. Il metodo di rendicontazione consiste nella presentazione di *case studies*, opportunamente selezionati dal programma di ricerca valutato, che riassumono e documentano un *context statement* (una dichiarazione sul focus, tipologia e direzione della ricerca) e un *impact statement* che risponde alle seguenti quattro domande.

1. Come la comunità di ricerca ha coinvolto gli utilizzatori finali con lo scopo di affrontare questioni sociali, economiche, ambientali e/o culturali?
2. Quali nuovi prodotti, politiche, normative, paradigmi, attitudini, prospettive, etc hanno adottato o implementato gli utilizzatori a seguito del loro coinvolgimento con la comunità di ricerca?
3. Quali sono i benefici sociali, economici, ambientali e/o culturali dei nuovi prodotti, politiche, normative, paradigmi, attitudini, prospettive, etc adottati dagli utilizzatori?
4. Qual è l'ordine di grandezza dei benefici sociali, economici, ambientali e/o culturali per gli utilizzatori a seguito dell'implementazione della ricerca?

In sintesi l'impatto viene misurato in funzione dell'intensità delle relazioni con gli utilizzatori finali, dell'adozione dei risultati da parte di questi ultimi, dei benefici ottenuti e della loro misura di effetto.

Una lista di indicatori viene fornita per facilitare la redazione dei resoconti e la valutazione da parte di panel di esperti.

La RAND nelle sue conclusioni suggerisce l'adozione e l'adattamento di questo modello, raccomandando la raccolta dai dipartimenti delle università di un numero definito di casi studio, redatti in forma narrativa e integrati con dati quantitativi e qualitativi. Il programma REF 2014 ha quindi accolto il suggerimento di includere la valutazione di impatto – incentrata sull'utilizzo dei risultati e sulle relazioni con gli utilizzatori – e di adottare il sistema dei casi studio, avviando la prima sperimentazione del programma. Quasi settemila report di casi studio sono stati raccolti e valutati da numerosi panel di esperti comprendenti un totale di 898 accademici e 259 utilizzatori dei risultati della ricerca (<http://impact.ref.ac.uk/CaseStudies/>). Tutto il materiale del programma, compresi i casi studio, sono accessibili su internet, fornendo un formidabile database – unico nel suo genere – di esperienze e interpretazioni di impatto.

Greenhalgh e Fahy hanno approfittato di questa risorsa per indagare il tipo di disegno di studio, la natura dell'impatto e i meccanismi di azione descritti, in un campione di 162 dei 1586 casi studio inviati al

Panel A, dedicato alla ricerca clinica e sanitaria. Non è sorprendente che la maggior parte degli studi presentati sia di tipo quantitativo, prevalentemente studi clinici randomizzati. I destinatari principali della ricerca vengono individuati nei responsabili di politiche sanitarie in 133/162 studi, nei clinici in 88/162 studi e nell'industria in 15/162 studi. Solamente 8 studi identificano i pazienti, o i cittadini in generale, come destinatari dei risultati della loro ricerca. Quasi tutti i casi analizzati descrivono un modello lineare di impatto, con più di 2/3 che sostengono di avere influenzato linee guida per la pratica clinica, mentre più della metà ritiene di avere influenzato le scelte di politica sanitaria a livello nazionale o internazionale. Meno comune è il coinvolgimento dei clinici e la documentazione di impatto sulla riduzione di morbilità o mortalità, mentre la ricerca commissionata dai decisori documenta un rapporto stretto con i committenti. Questo articolo verrà seguito da altri e numerosi studi che analizzeranno la miniera di informazioni raccolte dal REF 2014, che a sua volta intende perseguire e perfezionare la valutazione dell'impatto della ricerca e dei programmi di finanziamento.

Lo scopo di REF 2014 è stato – infatti – di sperimentare l'applicabilità, accettabilità ed efficacia del sistema, finalizzato sia a informare e supportare gli enti di finanziamento nel valutare le decisioni prese in passato e nell'orientare quelle future, sia a sostenere gli istituti di ricerca nel meglio programmare e focalizzare le proprie attività di ricerca. L'esercizio guidato per la rendicontazione ha indotto i ricercatori a riflettere sull'impatto del loro lavoro e a documentarlo, mentre la forma narrativa ha permesso ai ricercatori stessi di proporre o attribuire un significato proprio all'impatto della loro ricerca, fornendo un prezioso feedback al programma stesso e alle assunzioni in esso contenute.

L'articolo di commento conclude che la valutazione formale dell'impatto della ricerca è solo agli inizi. Queste sperimentazioni, così dettagliatamente rendicontate, possono tuttavia rappresentare un'importante fonte di ispirazione per chi voglia tentare esercizi simili a livello del proprio istituto (per esempio, Irccs) o a livello di ente di finanziamento (dalle fondazioni private ai programmi regionali o locali di assegnazione di fondi per la ricerca). Lo sviluppo e la condivisione di simili esperienze potrebbe far emergere una valida proposta – derivata da sperimentazioni fatte in Italia e in diversi ambiti delle nostre eccellenze – per un modello istituzionale di valutazione della ricerca nel nostro paese, in grado di informare e migliorare – e magari anche incrementare – i nostri sistemi di finanziamento.