

## Registri sanitari: aspetti metodologici

Paolo Bruzzi

**Riassunto.** I registri sanitari e gli studi osservazionali sono entità completamente differenti, e nonostante i registri siano spesso concepiti come studi osservazionali di coorte, o forniscano i dati per studi necessari per studi di coorte o caso-controllo, questo non sempre è vero: molti registri hanno scopi specifici che implicano requisiti metodologici completamente diversi da quelli di uno studio osservazionale. Naturalmente, se un registro viene concepito con lo scopo di condurre uno studio osservazionale, questo ne determinerà i metodi. È necessario anche sottolineare la differenza tra un registro e la raccolta e analisi di dati sanitari di routine: un registro può richiedere la raccolta di dati da fonti amministrative diverse, o la raccolta di dati ad hoc, non disponibili altrove, e deve garantire la completezza e la qualità delle informazioni che contiene. È stato ripetutamente suggerito che gli enormi archivi di informazioni sanitarie che si stanno accumulando presso tutti i sistemi di governo della sanità dovrebbero essere sfruttati direttamente per valutazioni di appropriatezza e di efficacia. Questo è molto pericoloso, perché i bias potenziali che già minacciano tutti gli studi osservazionali possono renderli privi di qualsiasi valore se i tradizionali standard metodologici non vengono seguiti rigorosamente. Analogamente, molta prudenza dovrebbe essere utilizzata nelle valutazioni di efficacia basate su studi osservazionali.

**Parole chiave.** Registri sanitari, studi osservazionali, valutazioni di efficacia.

### Introduzione e definizioni

Su cosa siano i registri (in Sanità) e su quale sia la loro funzione non c'è pieno accordo, o meglio c'è molta confusione. Per questo è importante partire da alcune definizioni, che, per quanto arbitrarie, permetteranno di capire a chi legge cosa intendiamo quando parliamo di registri.

Per la definizione di registro ci si rifaccia a quella fornita da Giuseppe Costa, per cui un registro è “un'attività continua e sistematica di rilevazione nel tempo...”.<sup>1</sup> Quindi un registro non è sinonimo di studio, e in particolare di studio osservazionale (tabella 1): uno studio nasce per rispondere a un quesito, mentre un registro non è identificato da uno specifico obiettivo conoscitivo ma può avere

*Health registries: methodological issues.*

**Summary.** Health registries and observational studies are different entities, and although several registries are conceived as observational cohort studies, or provide the data that are needed for cohort or case-control studies, this is not always true: many registries have specific aims that imply methodological standards that differ completely from those of an observational study. Of course, whenever a Registry is conceived with the aim of conducting an observational study, this will dictate its methods. The difference between a registry and the simple assembling and analysis of routine clinical data must be stressed as well: a registry may require that collection of data from different administrative sources, or of original data not available elsewhere, and must guarantee the quality and completeness of the information it contains. It has been repeatedly suggested that the huge databases of health information that are currently assembled by most health governing bodies should be exploited directly for evaluations of appropriateness and effectiveness. This is extremely dangerous, because the inherent potential biases already affecting all observational studies may completely undermine their validity if the currently accepted methodological standards are relaxed. Similarly, the use of observational studies for the assessment of the effectiveness of health technologies must be considered with caution.

**Key words.** Health registries, observational studies, evaluations of effectiveness.

molteplici obiettivi, come classicamente i registri di mortalità e i registri Tumori, ed è perfino possibile che la sua funzione primaria non sia di tipo conoscitivo: se uno pensa ai Registri AIFA non li vede primariamente come strumenti di conoscenza, quanto come iniziative volte a circoscrivere l'utilizzo di alcuni farmaci. Un altro esempio potrebbe essere quello di registri di patologia, utilizzabili come serbatoi di informazioni per sperimentazioni cliniche.

Quello che caratterizza il registro è invece l'attività di rilevazione sistematica e continua: rilevazione di cosa? Costa distingue “eventi o stati” relativi: a) allo stato di salute, b) a esposizioni a fattori di rischio, c) a trattamenti, d) a combinazioni delle informazioni precedenti. Per quanto un po' rigida,

| Registri   | Studi osservazionali  |
|--|---|
| Possono essere multiscopo  | Hanno uno/pochi obiettivi precisi di ricerca/valutazione                            |
| Non hanno un termine previsto  | Hanno un termine previsto   |
| Di solito includono principalmente dati raccolti per altri scopi (clinici, amministrativi) | Spesso prevedono dati raccolti ad hoc e talora "azioni" (es. questionari, prelievi) |
| Raccolgono molte informazioni che "potrebbero" rivelarsi utili                             | La raccolta dati è finalizzata agli obiettivi dello studio                          |
| Dovrebbero alimentare o facilitare studi osservazionali                                    | Prevedono un protocollo di ricerca con un piano statistico                          |
| Se hanno precisi obiettivi di ricerca, sono studi osservazionali                           |   |

questa precisazione è utile, perché ci permette di classificare i registri in base all'oggetto della rilevazione, e quindi in base alla loro vera natura.

Una seconda distinzione, ancora più importante, è quella tra registri e database amministrativi: cosa caratterizza un registro rispetto all'immane attività di raccolta dati che avviene routinariamente in Sanità? Possiamo identificare due differenze principali: a) nel registro non si parte dai dati disponibili ma dai dati necessari per gli scopi del registro: questi dati possono essere informazioni che sono raccolte *ad hoc*, e/o informazioni che devono essere estratte da basi dati diverse; b) la raccolta e l'archiviazione dei dati sono strutturate e sottoposte a controlli di qualità, in modo da rendere il patrimonio di informazioni raccolte nel registro adatto e disponibile per i suoi scopi.

L'importanza di questa seconda distinzione nasce dalla tendenza, diffusa tra molti informatici, di ritenere che la "quantità" di informazioni contenute nei database amministrativi sia tale da rendere le preoccupazioni degli epidemiologi superflue o addirittura antistoriche: da questi database sarebbe possibile estrarre, con procedimenti informatici, la maggior parte delle informazioni che si ottengono faticosamente con gli studi epidemiologici e clinici, persino quelle relative all'efficacia dei trattamenti. A questo proposito serve ricordare che l'incremento della dimensione di un campione è utile per ridurre l'errore campionario, ma è inutile o, anzi, dannoso in presenza di un errore sistematico (bias), perché conferisce alle stime ottenute una "precisione" che è del tutto ingiustificata e che può ingenerare false sicurezze. Analisi di dati amministrativi possono servire solo a generare ipotesi, mai a dimostrarle, in assenza di un disegno di studio organico e appropriato e di un'analisi statistica adeguata.

Nella progettazione di un registro, il fattore più importante è rappresentato dai suoi scopi, e non ha quindi molto senso parlare dei requisiti metodologici di un registro, se non in riferimento ai suoi scopi specifici, che sono quelli che determinano questi requisiti.

### Scopi di un registro e implicazioni metodologiche

Per identificare possibili scopi di un registro sanitario può essere utile fare riferimento alla classificazione proposta nella checklist presentata nel contributo "Riflessioni e confronti sui limiti e i vantaggi dei registri"<sup>2</sup> (tabella 2).

Visto che i requisiti metodologici di un registro dipendono largamente dai suoi scopi primari, è utile discutere separatamente i vari casi.

Tabella 2. *Possibili scopi di un registro sanitario.*

| Valutare   |
|--|
| a) la storia naturale della malattia   |
| b) la qualità dell'assistenza  |
| c) l'erogazione di servizi e il consumo di risorse   |
| a) l'efficacia e/o costo-efficacia degli interventi, la <i>comparative effectiveness</i> , e la sicurezza degli interventi |

REGISTRI COME SUPPORTO PER STUDI CHE SI PROPONGONO DI DESCRIVERE LA STORIA NATURALE DI UNA MALATTIA

Gli obiettivi di questi studi possono essere molteplici, non necessariamente in contrasto tra loro. Possono esserci motivi scientifici (studi biologici, ma anche la necessità di capire se esistono bisogni sanitari insoddisfatti). Ci può essere l'esigenza di fornire e ottenere corrette informazioni prognostiche, utili quando pazienti e medici devono stabilire l'iter terapeutico, e in questo caso è fondamentale l'informazione sui fattori che determinano la prognosi; possono esserci infine esigenze organizzative, per cui è utile sapere come dimensionare e dove collocare determinati servizi.

Sul piano metodologico, si tratta di pianificare dei classici studi di coorte, con i problemi metodologici che ne conseguono e che possono essere molto più complessi di quanto possa apparire. Tra questi problemi, quello della *rappresentatività* del-

la coorte è fondamentale, il più difficile da risolvere. I pazienti eleggibili dovrebbero essere identificati, inclusi nelle coorte e seguiti a partire dalla diagnosi, senza esclusione di alcun tipo. Questo criterio (la "consecutività" dei pazienti) non è però che un prerequisito, necessario per avviare una riflessione, molto più complessa, sulla possibilità di estrapolare la distribuzione di probabilità dei vari esiti osservata in questa coorte a pazienti o gruppi di pazienti futuri, perché le residue possibilità di bias sono tutt'altro che secondarie:

- I. innanzitutto, una coorte di pazienti reclutata presso un centro può essere inadeguatamente rappresentativa della totalità dei casi, per la possibilità che l'accesso al centro (ad esempio un centro specializzato) sia legato ad un profilo prognostico mediamente più o meno favorevole, a una particolare evoluzione della malattia (di solito negativa). In questo senso, le coorti che si basano su registri di popolazione sono molto più affidabili, se la qualità e la completezza della registrazione sono elevate, ma anche questo può non essere sufficiente.
- II. Le modalità di diagnosi prevalenti possono infatti condizionare pesantemente la storia naturale della malattia in una coorte di pazienti, per i ben noti problemi di *lead time* e di *length bias*: ad esempio, in un'area dove la pressione diagnostica è maggiore si osserverà una prognosi migliore, anche a parità di stadio. Le osservazioni nelle popolazioni sottoposte a screening confermano l'importanza di questo fenomeno per cui non si può parlare di storia naturale di una malattia se non in relazione alle modalità diagnostiche prevalenti nella popolazione studiata. Un caso clamoroso è quello del carcinoma prostatico: negli anni '80-90 l'incidenza di carcinoma prostatico era 4-5 volte superiore negli USA, dove l'utilizzo del PSA in soggetti asintomatici si era già diffuso, rispetto all'Italia, mentre la mortalità era simile nei due paesi. Con la diffusione del test anche in Italia, l'incidenza è salita fino ad avvicinare quella statunitense mentre le due mortalità hanno continuato ad essere simili.
- III. Il termine "storia naturale" è ormai inappropriato perché è quasi impossibile osservare popolazioni che non abbiano ricevuto terapie (con l'ovvia eccezione delle malattie per cui non esiste alcuna terapia efficace), per cui sarebbe più corretto parlare di storia della malattia sottoposta alle terapie standard. Qualsiasi inferenza a partire dai risultati di uno studio dovrebbe tenere conto di questo aspetto.
- IV. Le stesse problematiche (modalità diagnostiche e terapie), oltre a rendere problematica la definizione di storia naturale, comportano la *possibilità di bias*, perché spesso la coorte viene ricostruita retrospettivamente a partire da punti di "raccolta" di pazienti, vale a dire specifici esami diagnostici o particolari terapie.
- V. Un ultimo problema è quello relativo alla rilevazione dell'outcome, che pone quasi sempre

problemi di affidabilità e di completezza della rilevazione. Questo può non essere un problema in alcune malattie come i tumori, nei quali l'outcome di maggiore interesse è la sopravvivenza, ma lo è sicuramente nelle patologie a bassa letalità (ad esempio la sclerosi multipla e altre malattie neurologiche o le malattie reumatiche) dove l'interesse si concentra su altri esiti, come la disabilità, specifici sintomi, o altre dimensioni di difficile misurazione.

Per tutti questi motivi, un registro finalizzato a valutare l'evoluzione e la prognosi (con i fattori che la determinano) di una malattia deve essere concepito come uno studio epidemiologico e disegnato e condotto di conseguenza, con tutta la cura e gli accorgimenti necessari per garantirne la validità sia interna che esterna.

#### REGISTRI COME SUPPORTO PER STUDI FINALIZZATI A VALUTARE LA QUALITÀ DELL'ASSISTENZA

Qui siamo in un ambito quasi opposto al precedente, perché l'oggetto della nostra osservazione non è la malattia ma un Sistema Sanitario, oppure un particolare setting assistenziale (Ospedale, Reparto), e la sua gestione di uno specifico gruppo di pazienti. Anche in questo caso, è molto importante fare chiarezza: la qualità dell'assistenza può essere valutata direttamente, attraverso le classiche valutazioni di appropriatezza e di efficienza, nelle quali si confrontano i comportamenti e le performance nel contesto sotto osservazione con quanto prescritto da protocolli o linee guida riconosciute; oppure indirettamente, utilizzando indicatori surrogate, che dovrebbero permetterci di "stimare" la qualità dell'assistenza.

Partiamo dalle valutazioni "dirette" della qualità dell'assistenza, che pongono problemi di tre tipi:

- I. la scelta degli indicatori su cui effettuare le misurazioni; questi dovrebbero essere validati, e coprire tutte le principali aree di appropriatezza clinicamente rilevanti, dai test diagnostici, di stadiazione e follow-up, alle varie pratiche terapeutiche, agli aspetti procedurali o complementari, altrettanto rilevanti, come, ad esempio, la gestione degli effetti collaterali delle terapie o il supporto psicologico e logistico per i pazienti e i loro parenti; un problema nasce dal fatto che mentre le linee guida sono spesso sufficienti per la definizione degli indicatori che fanno riferimento al "cosa" deve (o non deve) essere fatto, lo sono molto più raramente rispetto al "come" deve essere fatto: ad esempio, la tempistica, il setting assistenziale, il coinvolgimento delle competenze necessarie (chirurgo specialistico? Radiologo addestrato alla procedura specifica? ecc.). Questo spesso obbliga a scelte del tutto arbitrarie.
- II. La raccolta delle informazioni su alcuni tra questi indicatori, quasi mai disponibili come

dati correnti, e spesso neppure recuperabili in automatico da archivi informatizzati.

- III. La tempestività dell'intero processo di identificazione dei casi eleggibili, raccolta delle informazioni sugli indicatori di appropriatezza e analisi e interpretazione dei risultati. È chiaro che quando questo processo richiede anni (come spesso accade), non se ne vede l'utilità, se non in termini di storia della medicina: a cosa può infatti servire la valutazione della proporzione di pazienti che in passato, in un certo reparto, sono stati trattati in accordo con le linee guida del tempo? È verosimile che oggi i medici di quel reparto siano cambiati, le linee guida siano state modificate e quello che era appropriato è oggi inappropriato e viceversa.

Le prime due problematiche (l'identificazione di indicatori validi e la difficoltà di raccogliere le informazioni necessarie) fanno sì che spesso si utilizzino, come surrogato delle valutazioni di qualità dell'assistenza, *valutazioni di esito*, sia a breve che a lungo termine (esempio: sopravvivenza a 30 giorni dall'intervento o a 5 anni dalla diagnosi). Non si può fare a meno di entrare nella discussione sull'appropriatezza di questa scelta (che pure, anche di recente, ha visto importanti iniziative sponsorizzate dall'AGENAS a livello nazionale con forti ricadute, perlomeno a livello mediatico): si ritiene questa scelta nella maggioranza dei casi superficiale, irresponsabile e frutto di una profonda incultura metodologica, epidemiologica e clinica<sup>a</sup>. Non si capisce infatti perché, in questo ambito, valutazioni basate su un indicatore surrogato non validato (outcome come surrogato della qualità dell'assistenza), in un contesto del tutto osservazionale, possano essere considerate scientificamente valide, senza quelle profonde riserve con cui verrebbero accolte se riguardassero l'efficacia di un trattamento. Per spiegarsi meglio, se si accetta come scientificamente valido il confronto fra la sopravvivenza tra due gruppi di pazienti diagnosticati in due aree diverse, e si attribuiscono le

differenze osservate alla differente qualità dell'assistenza, si sta di fatto affermando che il sistema assistenziale di un'area è più "efficace" di quello della seconda: se però un confronto di questo tipo è universalmente e giustamente considerato inaccettabile per dimostrare l'efficacia di un farmaco, non si capisce come possa essere accettato per dimostrare l'efficacia di "sistemi" assistenziali.

Il primo requisito di uno studio finalizzato alla valutazione dell'appropriatezza dei comportamenti, o più in generale della qualità dell'assistenza, è quindi l'utilizzo di indicatori di processo, diretti o surrogati, se validati. Il secondo requisito è la rappresentatività del campione di pazienti inclusi nello studio, il che esclude, tendenzialmente, i registri dove i pazienti sono inclusi sulla base dell'impiego di una certa tecnologia, terapeutica o diagnostica: questi registri possono essere utili per altri scopi (vedi oltre) ma non possono fornire un quadro credibile dei percorsi assistenziali seguiti dalla totalità dei pazienti affetti da una certa patologia, inclusi quelli che seguono percorsi del tutto anomali (che in certe patologie sono una quota anche importante). La rappresentatività può essere verificata esaminando i meccanismi di campionamento della coorte di pazienti su cui si valuta la qualità dell'assistenza. Solo nei casi in cui la coorte include tutti i pazienti eleggibili, o il campionamento segue qualche logica che garantisce la validità statistica del campione, la rappresentatività è (statisticamente) garantita, e le valutazioni di appropriatezza possono essere considerate statisticamente valide.

Un registro di patologia (di popolazione ma anche ospedaliero), in questi casi, è l'unico strumento che possa garantire questa rappresentatività, mentre, come già detto, si dovrebbe diffidare delle valutazioni basate su registri di procedure o di terapie. Certamente, rispetto a tutti i problemi appena discussi (scelta degli indicatori, raccolta di tutte le informazioni necessarie, tempestività), questo problema sembra di facile soluzione, ma non va trascurato; si pensi ad uno studio che si proponesse di valutare l'appropriatezza della gestione di una patologia cronica: nel caso del diabete e dei tumori la presenza dei "registri" di esenzione ticket può permetterci la ricostruzione di coorti abbastanza rappresentative, ma se passiamo ad altre patologie (es. artrosi, dolore cronico, emicrania, ecc.), è quasi scontato che qualsiasi "registro" includerà preferenzialmente pazienti trattati adeguatamente o sovratrattati mentre i casi di sottotrattamento saranno poco rappresentati.

In sostanza, l'utilizzo di registri per valutazioni sulla qualità dell'assistenza dovrebbe essere uno dei campi di applicazione più rilevanti dei registri di patologia; gli studi in questo settore sono però penalizzati dal fatto che le difficoltà che si incontrano se si vuole assicurare a questi studi la necessaria validità scientifica, che sono oggettive e difficili da superare, sono spesso affrontate con una certa sciattezza metodologica.

<sup>a</sup> Per confermarlo sarebbe sufficiente considerare quanto spesso si osservano, tra popolazioni di aree diverse, differenze anche grandi nella sopravvivenza di tumori per i quali le terapie disponibili hanno scarsa o nulla efficacia, e non esiste, né sarebbe di alcuna utilità, un'attività di prevenzione secondaria; è il caso, ad esempio, di tumori come il mesotelioma o il carcinoma pancreatico. Se queste differenze non possono essere attribuibili a una diversa qualità dell'assistenza ma sono necessariamente dovute a uno o più bias (alcuni dei quali ben noti), non si capisce perché, quando osservate in tumori dove esistono terapie efficaci, possano essere attribuite a una migliore qualità dell'assistenza. Questo errore, nel caso dei Registri Tumori, potrebbe nascere dal desiderio di promuoverne l'utilità come strumenti di sanità pubblica, utilità che resta tutta da dimostrare, anche in ambito preventivo. Con questo non voglio mettere in discussione il grande valore dei Registri Tumori in ambito puramente scientifico, ma sottolineare come la rilevanza che è stata data, sia in ambito scientifico che mediatico a queste valutazioni prive di alcuna validità, è pericolosa perché porta a sopravvalutare la qualità dell'assistenza in paesi, l'Italia, dove l'utilizzo dei test diagnostici e la qualità delle diagnosi sono meno rigorosi.

#### REGISTRI PER VALUTARE EROGAZIONE DI SERVIZI E CONSUMO DI RISORSE

Queste valutazioni rappresentano in realtà un'articolazione/estensione di quello precedente (valutazione della qualità dell'assistenza) perché non sono concepibili, e tanto meno interpretabili, se non in riferimento all'appropriatezza di utilizzo dei servizi erogati e delle risorse consumate. Infatti, solo se ci si trova nell'ambito di un impiego appropriato (vale a dire di strumenti utilizzati nel contesto in cui la loro efficacia è provata), si potrà parlare di bisogni più o meno soddisfatti, e di servizi e risorse più o meno adeguati ai bisogni. Un ambito di confine è quello relativo all'utilizzo inappropriato di strumenti organizzativi (esempio ricovero ospedaliero) dove il monitoraggio continuo attraverso attività di registrazione può rivelarsi molto utile. In questi ambiti le problematiche metodologiche non sono molto diverse da quelle che si incontrano negli studi sulla qualità dell'assistenza, semplificate dal fatto che le valutazioni sono principalmente di tipo quantitativo.

#### REGISTRI PER VALUTARE EFFICACIA, E COSTO/EFFICACIA DEGLI INTERVENTI – COMPARATIVE EFFECTIVENESS RESEARCH

È questo l'ambito dove c'è maggiore bisogno di chiarezza, perché è quello in cui ambiguità terminologica e concettuale si rinforzano reciprocamente in maniera più drammatica. Per tale ragione il resto di questa relazione sarà dedicato a questo tema.

#### **Registri/studi osservazionali e valutazioni di efficacia**

I termini di questo problema devono essere affrontati con ordine.

- a) Il tema è quello delle valutazioni di efficacia nell'ambito di studi osservazionali. Questi studi possono nascere come registri, o possono essere innestati su registri preesistenti o di nuova istituzione, ma si differenziano dai registri per alcuni aspetti critici:
  - i. nascono con un preciso obiettivo primario (valutazione di efficacia).
  - ii. La loro metodologia e le informazioni raccolte nel loro ambito sono esplicitamente finalizzate al raggiungimento di questo obiettivo (e di quelli secondari), e possono prevedere azioni (es. raccolta dati, ma anche esami) che esulano dalla normale pratica clinica, pur non alterandone gli aspetti sostanziali (terapie).
  - iii. Dovendo raggiungere questo obiettivo, hanno un orizzonte temporale e dimensioni predefiniti e comunque limitati al raggiungimento dell'obiettivo.
  - iv. Tutti questi aspetti sono chiaramente definiti in un protocollo di ricerca preparato preventivamente.

Ovviamente, un registro può nascere con tali caratteristiche, ma in questo caso si tratta di un vero e proprio studio osservazionale prospettico, e come tale deve essere definito e caratterizzato. Viceversa, uno studio osservazionale di efficacia retrospettivo (= di coorte retrospettivo) può essere innestato su un registro preesistente, con tutti i problemi metodologici che questo può comportare.

- b) Nell'ambito delle valutazioni di efficacia, e degli studi ad esse finalizzati, il problema principale è quello della Validità Interna (o Validità Statistica), ovvero del controllo dei bias (distorsioni) che possono produrre errori nella stima dei parametri che si vogliono valutare, in questo caso l'efficacia di un intervento. Ricordiamoci che qualsiasi discussione sulla Validità Esterna (=implicazioni) di uno studio è priva di senso se lo studio presenta distorsioni tali da minarne la Validità Interna. Lo strumento principe per ottenere stime di efficacia statisticamente valide è il trial randomizzato (RCT). Lo stesso RCT non garantisce l'assenza di difetti che possono portare a bias, per cui, prima di valutare le implicazioni cliniche dei suoi risultati, è sempre necessario sottoporlo a un vaglio metodologico critico. Nell'ambito di un RCT, però, queste valutazioni sono interamente concentrate sulle procedure, che, se progettate e condotte in accordo con gli standard riconosciuti, garantiscono l'assenza di bias, e in particolare l'assenza di un bias di selezione dei soggetti assegnati al trattamento sperimentale rispetto a quelli assegnati al trattamento standard. Questa garanzia non è mai possibile al di fuori del RCT, per cui negli studi non randomizzati si deve sempre presumere che qualsiasi valutazione di efficacia possa essere "sporcata" da qualcuno tra i molti possibili bias che affliggono questi studi. Quando un gruppo di ricercatori che hanno disegnato e condotto uno studio senza un gruppo di controllo randomizzato vuole utilizzarne i risultati per dimostrare l'efficacia di un trattamento, deve quindi portare tutti gli elementi conoscitivi necessari per sostenere la validità dei suoi risultati, e per rigettare l'ipotesi che siano frutto di uno di questi possibili bias.
- c) È importante distinguere gli studi osservazionali dalle sperimentazioni cliniche senza gruppo di controllo randomizzato, anche se i due tipi di studi hanno in comune vari tipi di possibili bias. Una sperimentazione clinica si pone come obiettivo la valutazione di efficacia e quindi dovrebbe valutare tecnologie o interventi la cui efficacia non è ancora provata (non utilizzabili nella pratica clinica) o confrontare gli interventi di efficacia nota, per stabilire uno standard di comportamento. L'utilizzo della tecnologia/intervento nella sperimentazione dovrebbe quindi essere finalizzato alla valutazione di efficacia.

Nello studio osservazionale l'utilizzo della tecnologia/intervento ha invece finalità cliniche, e la valutazione di efficacia si innesta su questo utilizzo come obiettivo di ricerca e riguarda solo tecnologie già entrate nella pratica clinica (anche se sulla base di evidenze di efficacia spesso inadeguate).

Storicamente, le sperimentazioni senza gruppo di controllo randomizzato avevano tre possibili motivazioni:

- i. studi di sicurezza o attività (Fase I o II) che non richiedevano un gruppo di controllo per la specificità del fenomeno studiato (es. leucopenia, massiva risposta tumorale).
- ii. Insipienza metodologica degli investigatori, che ignoravano le regole fondamentali della ricerca clinica e/o malafede degli sponsor, con la connivenza delle autorità regolatorie (vedi gli scandali italiani sul farmaco dei primi anni '90<sup>b</sup>).
- iii. Motivazioni etiche, nel caso di malattie a cattiva prognosi dove non esistevano terapie efficaci e il trattamento sperimentale appariva promettente.

Di queste tre motivazioni, la prima è sempre valida, a condizione di riconoscere esplicitamente che la finalità di uno studio di Fase I o II non è quella di valutare l'efficacia. La seconda è (quasi) completamente sparita, mentre gli spazi per la terza si sono drasticamente ridotti, con il riconoscimento che molto spesso la presunta efficacia di un nuovo trattamento non viene confermata nel RCT. Questo però vale soprattutto per i farmaci, dove i vincoli regolatori sono molto più rigorosi, molto meno per il restante universo degli interventi e tecnologie di vario tipo, dagli interventi chirurgici ai test diagnostici, dai programmi di prevenzione alle tecnologie, ad esempio computerizzate o robotizzate, che spesso sono sottoposte al vaglio di sperimentazioni cliniche largamente inadeguate sul piano metodologico, per poi essere introdotte nella pratica clinica. In questa situazione di conoscenze insufficienti si crea uno spazio per studi "osservazionali" che cerchino di valutare l'efficacia della tecnologia.

- d) L'esigenza di studi osservazionali finalizzati a valutare l'efficacia di una tecnologia viene spesso giustificata, oltre che da insufficienti conoscenze sulla sua efficacia "assoluta" (rispetto a nessun trattamento) o "relativa" (rispetto ad altri trattamenti di provata efficacia) dalla volontà di acquisire informazioni sulla sua "real world effectiveness" (efficacia nelle condizioni di utilizzo routinario) in contrasto con la sua "efficacy" (efficacia nelle condizioni di utilizzo ottimali che si realizzano nel clinical trial). Senza entrare nel

merito della problematica relativa alla distinzione tra *efficacy* ed *effectiveness*, si vorrebbe contestare l'ipotesi, sostenuta apoditticamente da molti, che sia possibile valutare *l'effectiveness* al di fuori di un contesto randomizzato. Perché per *l'effectiveness* non dovrebbero valere tutte quelle riserve metodologiche che rendono invece inaffidabile tale valutazione per *l'efficacy*?

- e) Le sperimentazioni cliniche (con o senza gruppo di controllo randomizzato) sono per definizione prospettiche. Gli studi osservazionali possono essere prospettici o retrospettivi, o avere una componente prospettica e una componente retrospettiva. Una discussione approfondita di tutti i possibili bias che possono minare la validità degli studi (sperimentali o osservazionali) senza gruppo di controllo randomizzato esula dagli scopi di questa discussione. Negli studi osservazionali i possibili bias sono però più difficili da controllare:
  - i. Nella sperimentazione non controllata la decisione di inserire un paziente (e quindi di sottoporlo al trattamento sperimentale) è regolata da rigidi criteri e procedure, che, in teoria, potrebbero permettere di identificare un gruppo di pazienti con caratteristiche simili ma non trattati, da utilizzare come gruppo di controllo. Nello studio osservazionale, invece, l'utilizzo del trattamento è in qualche modo arbitrario, e quindi condizionato da molti fattori difficilmente quantificabili (es. stadio e prognosi, comorbidità, preferenze del paziente) soppesati in modo diverso da ogni medico, il che rende molto difficile l'identificazione di un gruppo di controllo in qualche modo confrontabile.
  - ii. Nello studio osservazionale il gruppo di controllo può essere selezionato tra pazienti della stessa coorte (diagnosticati nello stesso periodo e centro, ma non sottoposti al trattamento in studio) oppure tra pazienti diagnosticati in un altro periodo e/o centro. La prima soluzione ha il vantaggio di garantire una sostanziale omogeneità nelle procedure di diagnosi, stadiazione e soprattutto follow-up dei pazienti, ma lo svantaggio che trattati e non trattati differiranno, mediamente, nei fattori che hanno portato alla decisione di utilizzare o meno il trattamento in studio. La seconda soluzione potrebbe in teoria consentire di evitare questo bias di selezione, ma la confrontabilità dei due gruppi in termini di criteri di diagnosi e stadiazione e in termini di rilevazione di esito e di tossicità diventa molto più problematica. Per ovviare a questi bias sono state proposte varie tecniche statistiche che dimostrate finora abbastanza inefficienti.
  - iii. Lo studio osservazionale (come pure la sperimentazione non controllata) è per definizione in aperto (senza procedure di mascheramento come il doppio cieco ottenuto grazie al placebo), con tutti i bias che ne possono derivare, specie nella valutazione degli endpoint.

<sup>b</sup> Scandali farmaci in Italia anni '90: qualche cenno. [http://archiviostorico.corriere.it/1994/luglio/19/Lorenzo\\_Poggiolini\\_145\\_capi\\_accusa\\_co\\_0\\_9407191813.shtml](http://archiviostorico.corriere.it/1994/luglio/19/Lorenzo_Poggiolini_145_capi_accusa_co_0_9407191813.shtml)

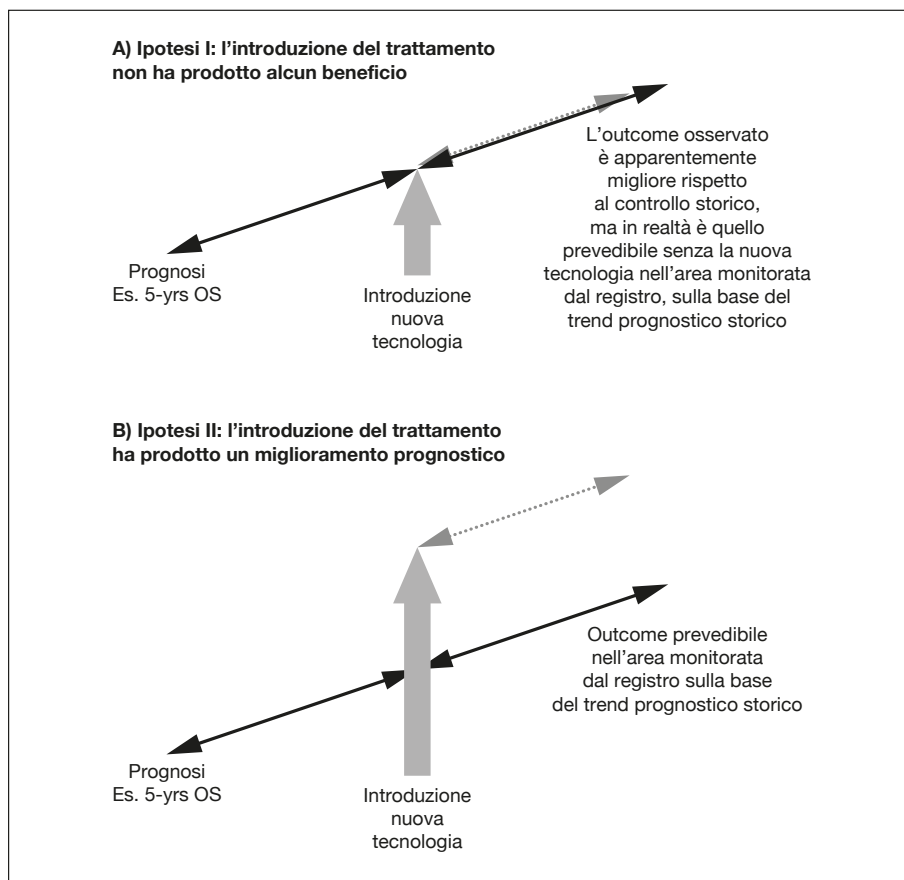


Figura 1. Registri di patologia per fornire controlli storici agli studi di efficacia.

Con tutte queste riserve, gli studi osservazionali, e in particolare quelli derivati da registri, possono avere un ruolo importante nella valutazione dell'efficacia delle tecnologie.

- f) Uno studio osservazionale, anche retrospettivo, può e dovrebbe essere disegnato e condotto sul modello di un trial prospettico randomizzato: questo significa definire la popolazione da cui è estratto il campione in studio, utilizzare criteri di selezione chiari, ricostruire la coorte di pazienti eleggibili che “sarebbero dovuti” entrare in studio e adottare nelle analisi un approccio simile alla cosiddetta “intention to treat”, per poter identificare i pazienti persi al follow-up e i non-compliers. In questo modo è possibile evidenziare molti tra i possibili bias di uno studio osservazionale e valutarne la possibile rilevanza. Nel disegno dello stesso studio, è indispensabile identificare preventivamente il gruppo di controllo in modo da massimizzare la possibilità di confrontare i due gruppi.
- g) Uno studio osservazionale basato sui dati di un registro di patologia può fornire un “gruppo di controllo storico” affidabile per una sperimentazione clinica che, per qualche motivo, non includesse un gruppo di controllo randomizzato. Questo può essere particolarmente importante nel caso di patologie molto rare, dove non sono possibili sperimentazioni cliniche canoniche in termini di dimensioni ma spesso anche sul pia-

no metodologico. In particolare, un registro di popolazione mette a disposizione un campione di pazienti non selezionati dal quale può essere possibile estrarre il sottogruppo di pazienti che rispondono ai criteri di eleggibilità dello studio, fornendo stime sull'outcome della malattia “atteso” senza l'utilizzo della terapia in studio.

- h) Un possibile problema in questo tipo di confronti nasce dal fatto che la prognosi della maggior parte delle malattie non è costante nel tempo, ma tende a mostrare un lento ma continuo miglioramento, a prescindere dall'introduzione di nuovi trattamenti. La disponibilità di dati di registro su tempi medio-lunghi può essere molto utile per superare questo problema: infatti, grazie alla ricostruzione dell'evoluzione nel tempo della prognosi di una patologia, può essere possibile valutare, nell'ambito di uno studio osservazionale, se l'introduzione di una nuova tecnologia si è associata con una modificazione drastica di questo trend storico, attribuibile a questa introduzione (figura 1)<sup>c</sup>.

<sup>c</sup> I registri di patologia, se rappresentativi di tutti i casi incidenti in una determinata popolazione, possono essere molto utili per valutare se l'introduzione di una nuova tecnologia si è tradotta in un cambiamento importante negli outcome (prognosi e tossicità) della patologia interessata dalla tecnologia, permettendo di valutarne il trend storico di evoluzione e quindi di valutare meglio gli effetti di questa introduzione.

- i) La disponibilità di dati affidabili da un registro di popolazione è indispensabile per evitare il bias di selezione, perché permette di considerare nello studio tutti i casi diagnosticati prima dell'introduzione della nuova tecnologia e tutti i casi diagnosticati dopo, compresi quelli che non hanno fatto uso della tecnologia.
- j) Infine, gli studi osservazionali sono indispensabili per acquisire maggiori informazioni sulla tossicità di un trattamento in specifici sottogruppi di pazienti (anziani, con comorbidità), per valutare le eventuali interazioni con altri farmaci, e per rilevare la presenza di effetti avversi non troppo frequenti. Considerando che la maggior parte delle sperimentazioni di Fase III arruolano, solitamente, alcune centinaia di pazienti (di cui solo la metà assume il farmaco sperimentale), in condizioni di salute mediamente superiori a quelle della totalità dei pazienti, è facile che eventi avversi, anche molto gravi, che hanno una frequenza inferiore a 1%, non si verifichino o non siano attribuiti al trattamento. Quando però il trattamento entra in commercio, e viene utilizzato da migliaia o decine di migliaia di pazienti, questi effetti possono diventare un problema di sanità pubblica rilevante, per cui è fondamentale la loro rilevazione tempestiva. Non sempre questi effetti avversi hanno una specificità tale da permettere l'attribuzione univoca al trattamento, come fu, ad esempio, nel caso della focomelia da talidomide<sup>d</sup>.
- k) La loro valutazione richiede quindi studi mirati che confrontino la loro incidenza in coorti di pazienti trattati e non trattati. La presenza di registri di patologia di buona qualità e di dimensioni importanti può sicuramente facilitare questi studi e promuoverne una maggiore attendibilità.

---

<sup>d</sup> La talidomide è un farmaco altamente teratogeno che fu venduto negli anni Cinquanta e Sessanta come tranquillante/sedativo da banco. Furono necessari diversi anni prima che le gravi malformazioni ai feti di madri che avevano assunto il farmaco in gravidanza fossero attribuite alla talidomide.

## Conclusioni

Oggi l'organizzazione e conduzione di un registro sanitario sono enormemente più facili e meno costose di 20-30 anni fa, grazie alle reti e agli archivi informatici di dati sanitari. Questo fatto deve però semmai indurre ad una maggiore tensione e vigilanza metodologica da parte degli addetti ai lavori, perché oggi qualsiasi soggetto che abbia un PC e l'accesso a questi archivi può produrre analisi inopportune e scorrette. Le conseguenze di simili "impresе" a livello mediatico, politico e anche organizzativo non vanno minimizzate, specie quando il soggetto di cui sopra riveste qualche posizione nel Servizio Sanitario Nazionale o in un'agenzia regionale o nazionale. È quindi fondamentale continuare a diffondere il corretto verbo metodologico e, in particolare, la distinzione tra le fonti di informazioni e gli studi che le utilizzano: per questi ultimi, valgono sempre le stesse coordinate interpretative, metodologiche e biostatistiche, codificate da decenni in sterminati manuali, articoli e corsi.

I registri, intesi come sistemi organizzati di raccolta dati finalizzati a uno o più scopi, hanno molto in comune con gli studi osservazionali, e quindi devono rispondere a requisiti molto simili, o perlomeno essere progettati e organizzati in modo da rendere possibile il loro utilizzo nell'ambito di studi che rispondono a questi requisiti. Altrimenti, rischiano di diventare uno strumento burocratico fine a se stesso, o, peggio, una fonte di informazioni autoreferenziali e spesso fuorvianti.

## Bibliografia

1. Costa G. Tassonomia e definizione dei registri clinici. *Recent Prog Med* 2015; 106: ...
2. Addis A, Costa E, De Palma R et al. Riflessioni e confronti sui limiti e i vantaggi dei registri. *Recent Prog Med* 2015; 106: 425-35.