

# L'importanza dell'integrazione di database amministrativi come fonte di dati epidemiologici: l'esperienza dell'Osservatorio ARNO

Editoriale

Marisa De Rosa<sup>1</sup>, Elisa Rossi<sup>1</sup>

Nella pratica clinica il medico è chiamato quotidianamente a effettuare scelte terapeutiche determinanti per la salute del paziente. Una corretta informazione e la disponibilità di dati epidemiologici diventano uno strumento importante per le decisioni cliniche.

Per meglio comprendere l'applicazione dell'epidemiologia nel campo della medicina generale, si riportano sinteticamente alcuni principi cardine dell'epidemiologia.

Riprendendo una definizione classica, l'**epidemiologia** è la disciplina che studia la **distribuzione** e le **cause** della **frequenza** delle malattie nelle popolazioni [1]. Questa definizione implica lo studio di tre componenti correlate tra loro:

- **frequenza:** quantifica la presenza e l'occorrenza della malattia;
- **distribuzione:** fornisce caratteristiche sulla malattia in studio nel tempo e nello spazio (chi, dove, quando), includendo anche i confronti con popolazioni o periodi storici differenti;
- **cause:** studia le cause legate all'evento in studio ed è fondamentale per descrivere i pattern della malattia e per poter formulare ipotesi su possibili cause o fattori di prevenzione [2].

Si divide fundamentalmente in due sezioni: descrittiva e analitica.

- **Epidemiologia descrittiva.** Viene utilizzata per identificare un fenomeno, descriverlo, misurarne la frequenza e/o studiarne lo sviluppo e la distribuzione nella popolazione, senza studiare le relazioni di causa-effetto. Si avvale di indicatori quali, tra i più usati, prevalenza e incidenza:

- la **prevalenza** misura il numero di individui di una popolazione che, in un dato momento, presentano la malattia. È utile nella programmazione sanitaria, in quanto misura l'impatto e la penetrazione che una malattia ha in un determinato territorio:

$$\text{prevalenza} \rightarrow \frac{\text{n. totale casi di malattia}}{\text{esposti a rischio}}$$

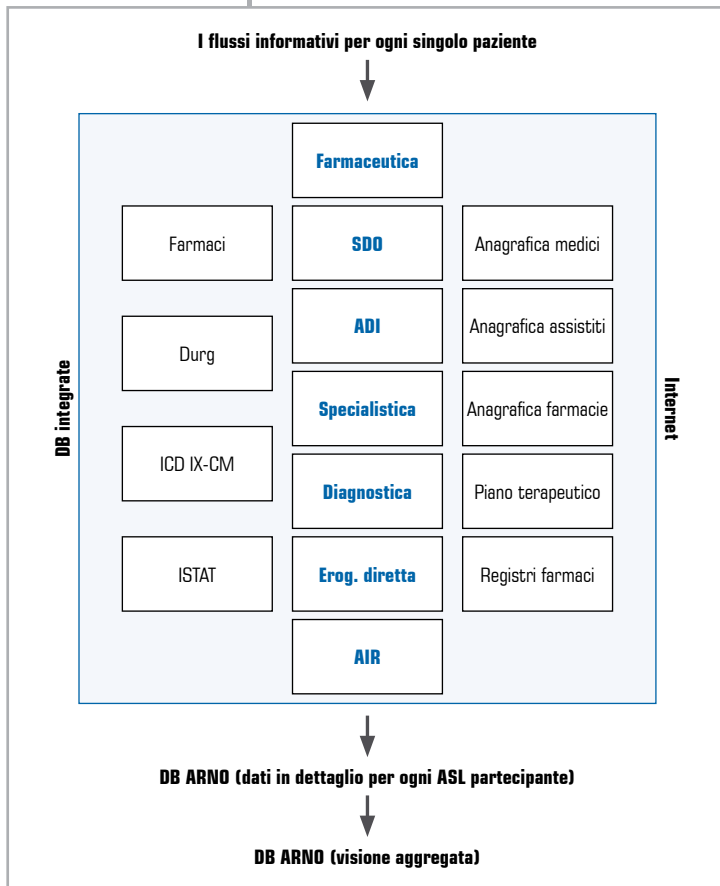
- l'**incidenza** misura la frequenza con cui si sviluppano nuovi casi di malattia in una popolazione precedentemente sana. Esprime la velocità di spostamento dallo stato di salute allo stato di malattia:

$$\text{incidenza} \rightarrow \frac{\text{nuovi casi di malattia in un periodo}}{\text{pop. (inizialmente sana) a rischio nel periodo}}$$

- **Epidemiologia analitica.** Viene usata, con la descrittiva, per identificare relazioni di causa-effetto tra fattori di rischio e malattia in una determinata popolazione e in un determinato periodo di tempo. Si avvale di misure di associazione (*odds ratio*, *hazard ratio*) e di metodi per valutare i fattori di rischio delle malattie: studi di coorte, studi caso-controllo.
- **Studi di coorte:** si effettuano per valutare i possibili fattori di rischio in un periodo di tempo. I gruppi sono selezionati in base all'esposizione o meno del fattore in esame (es. un farmaco) e si seguono nel tempo (follow-up) valutando l'evento in studio e la frequenza con cui si sviluppa. La durata dell'osservazione e la numerosità del campione dipendono dall'evento in studio.

<sup>1</sup> Dipartimento Sistemi Informativi e Servizi per la Sanità, CINECA

Corresponding author  
Dott.ssa Marisa De Rosa  
m.derosa@cineca.it



**Figura 1**  
Le banche dati integrate nell'Osservatorio ARNO

- **Studi caso-controllo:** metodologia usata per rafforzare lo studio degli eventi dei pazienti esposti al fattore in studio (casi) confrontati con un gruppo di pazienti non esposti al fattore in studio (controlli) [3]. La selezione del gruppo di controllo è un punto critico ai fini della validità dello studio perché questo gruppo deve essere il più possibile simile ai casi, soprattutto per

Popolazione totale (M = 48%; F = 52%; indice di vecchiaia = 147,7)	9.993.002
Numero medici di base	7.928
Numero medici pediatri	1.180
Numero di trattati	6.900.375
Numero ricette/anno	98.748.046
Banca dati ARNO storica	640 milioni di ricette

**Tabella I**  
I numeri dell'Osservatorio ARNO. Dati di 30 ASL di 7 Regioni diverse (aggiornamento settembre 2007)

quanto riguarda le caratteristiche della popolazione che potrebbero introdurre fattori di distorsione (es. si selezionano in genere due controlli per ogni caso con stesso sesso, età, zona geografica, esposizione ad altri fattori di rischio per la medesima patologia).

I database amministrativi, come gli archivi di prescrizioni farmaceutiche, sono caratterizzati da una grande disponibilità di dati con aggiornamento periodico e sono visti come **un'importante risorsa per condurre studi epidemiologici** nell'ambito del Servizio Sanitario Nazionale [4].

I consumi farmaceutici, che si ottengono dagli archivi delle prescrizioni, sono da tempo considerati come descrittori affidabili dell'esposizione dei pazienti ai farmaci come traccianti di patologia.

Tuttavia, per poter passare alla conduzione di veri e propri studi di farmacoepidemiologia è importante che questi archivi siano integrati sia con i dati anagrafici di popolazione che con altri descrittori di prestazioni sanitarie (es. ricoveri ospedalieri, prestazioni specialistiche).

In questo modo si può passare da una visione puramente descrittiva a una visione epidemiologica in cui **il paziente è al centro dello studio** con la possibilità di individuare popolazioni di riferimento. Queste popolazioni possono essere riferite a singole ASL, singoli distretti o anche a singoli medici, ma devono sempre essere pensate come facenti parte di gruppi **omogenei** che possono essere confrontati tra loro mediante **indicatori epidemiologici standardizzati** [5].

L'indicatore più efficace nella farmacoepidemiologia è l'indice di prevalenza (es. prevalenza d'uso di un farmaco), dove il numeratore è rappresentato dal **"numero di trattati"** che rappresenta il numero effettivo di persone in trattamento con un farmaco o affette da una determinata patologia.

Per misurare il consumo dei farmaci, in database che non dispongono del numero di trattati, l'indicatore maggiormente utilizzato è rappresentato dalle DDD/1.000 ab. die. La DDD (*Defined Daily Dose*: dose giornaliera definita) rappresenta la dose media giornaliera di un farmaco definita per la sua indicazione terapeutica principale. Questo indicatore è molto utile anche per effettuare confronti a livello internazionale.

Punto essenziale nell'uso di banche dati integrate di grandi dimensioni, con frequenti aggiornamenti dei dati provenienti da varie fonti, è la qualità dei dati. I dati devono infatti seguire rigorosi **controlli di qualità** per assicurare un alto standard qualitativo degli indicatori ottenuti per evitare sottostime nei dati.

- Gli errori derivano principalmente da:
- incompletezza delle informazioni (es. data di nascita mancante);

- codici errati nei processi di scrittura e lettura del dato (es. codice paziente errato).

Un esempio di database di tipo amministrativo come fonte di dati epidemiologici è l'Osservatorio ARNO. Attivo da più di vent'anni, è un osservatorio multicentrico delle prestazioni sanitarie erogate dal SSN al singolo cittadino, il cui scopo è quello di fornire, alle ASL convenzionate, un *Data Warehouse* clinico con i dati raccolti **per ogni singolo paziente** (es. ricette di prescrizione farmaceutica, schede di dimissione ospedaliera, specialistica ambulatoriale, ecc...).

Queste informazioni vengono integrate con i dati anagrafici e con ulteriori flussi informativi (dati socio/demografici), fornendo la possibilità di esplorare in dettaglio gli interventi sul singolo assistito e di elaborare progetti di *Disease Management*, *Bench Marking*, appropriatezza prescrittiva, studi di coorte e registri di patologia (Figura 1).

Ad oggi, l'Osservatorio registra le prestazioni sanitarie di una popolazione di 10 milioni di abitanti appartenenti a 30 Aziende Sanitarie Locali di 7 Regioni italiane (Tabella I).

Disponibile via web, mette a disposizione dei diversi profili di accesso (Direzione Generale, Distretto, medico) percorsi di analisi navigabili strutturati a seconda dei diversi profili di popolazione.

Le peculiarità dell'Osservatorio ARNO possono essere così riassunte:

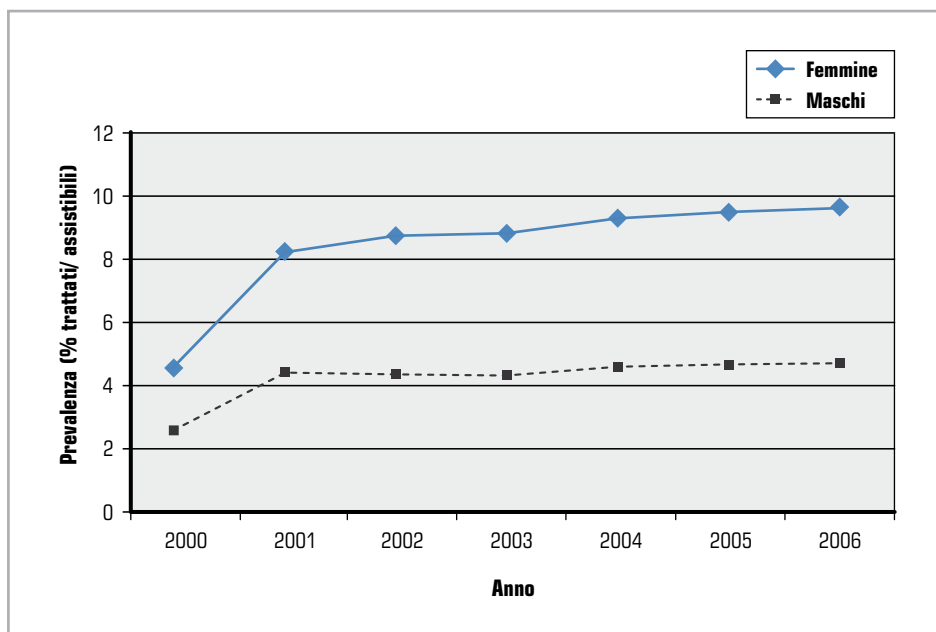
- è un database orientato al paziente (il paziente è al centro dello studio);

- raccoglie dati da più zone geografiche e consente la condivisione di dati omogenei utilizzando metodologie standardizzate;
- consente di valutare nel tempo la continuità e l'efficacia degli interventi sulla qualità e sui costi dell'assistenza;
- evidenzia popolazioni (bambini, donne, anziani, diabetici, ecc...) da seguire, sorvegliare e coinvolgere nel contesto di programmi di intervento;
- è uno strumento per valutare i bisogni terapeutici e l'appropriatezza delle risorse;
- misura l'impatto di nuovi farmaci;
- permette di effettuare studi di farmacovigilanza attiva;
- consente di valutare, in accordo con gruppi di medici e/o le loro rappresentanze, i problemi presenti nella pratica dei singoli (omogeneità/eterogeneità, carichi assistenziali, strategie terapeutiche, linee guida, audit clinico, appropriatezza, ecc ...) [6].

Nella Figura 2 e nella Tabella II sono descritti due esempi di studi epidemiologici su particolari coorti di pazienti, elaborati dai dati dell'Osservatorio ARNO.

In conclusione:

- i database clinici di popolazione indicano la strada del futuro e offrono l'opportunità di rendere la ricerca più vicina alla pratica quotidiana;
- la costituzione di popolazioni da evidenziare, seguire, sorvegliare è necessaria per valutare l'impatto degli interventi e l'applicabilità delle linee guida;



**Figura 2**

*Trend del consumo di antidepressivi nella popolazione. Fonte: Osservatorio ARNO-CINECA*

Nota: L'aumento dei consumi di antidepressivi dal 2000 al 2001 è imputabile principalmente al passaggio alla piena rimborsabilità degli SSRI

Caratteristiche descrittive	Pazienti con diabete (casi) *	Pazienti senza diabete (controlli) **	Var % diabetici vs non diabetici
N. pazienti	311.979	623.958	-
Prevalenza (min-max)	4,5% (3,0-5,3%)	-	-
Età media	67,6	67,6	-
% femmine	50,0	50,0	-
N. medio di ricette prescritte /anno	30,5	17,2	+ 77%
• per farmaci antidiabetici	7,2	-	-
• per altri farmaci	23,3	17,2	+ 36%
N. medio di ricoveri /anno	1,7	1,5	+ 10%
N. medio di prestazioni specialistiche/anno	30,5	24,6	+ 24%

**Tabella II**

*Osservatorio ARNO Diabete - metodologia caso-controllo con integrazione dei diversi flussi informati (farmaceutica territoriale, ricoveri ospedalieri, specialistica ambulatoriale) [7]*

\* I casi sono rappresentati dai pazienti in trattamento con un farmaco antidiabetico nel 2006

\*\* I controlli (2 per ogni caso) sono costituiti dai pazienti che non hanno ricevuto prescrizioni di antidiabetici e sono stati selezionati sulla base dello stesso sesso, età e medico curante dei casi

- la flessibilità e le possibilità offerte dalle nuove tecnologie possono essere sfruttate per sperimentare percorsi logici e operativi così da permettere di verificare attentamente “il guadagno conoscitivo” che si produce;
- è possibile oggi implementare e mantenere database a livello locale con possibilità di visione aggregata a livello centrale per

il confronto e lo sviluppo di metodologie comuni;

- l'approccio globale basato sull'integrazione di dati di popolazione, clinici ed economici, è finalizzato all'individuazione di un modello dell'intero processo che porti all'aumento del livello qualitativo dei servizi erogati e all'ottimizzazione dei costi complessivi.

**BIBLIOGRAFIA**

1. MacMahon B, Pugh TF. Epidemiology: principles and methods. Boston: Little Brown, 1970
2. Hennekens CH, Buring JE. Epidemiology in medicine. Boston: Little Brown, 1987
3. Armitage P, Berry G. Statistica medica. Milano: McGraw-Hill, 1996
4. Monte S, Fanizza C, Romero M, Rossi E, De Rosa M, Tognoni G. Database amministrativi come risorsa accessibile e strumento efficiente per l'epidemiologia cardiovascolare. *Giornale Italiano di Cardiologia* 2006; 7: 206-16
5. Tognoni G. Introduzione al progetto ARNO, rapporto 1996-1997. Roma: Cineca, 1997
6. De Rosa M (a cura di). L'utilizzo dei dati ARNO nella rete delle ASL. Documentazione di 20 anni di attività. Rapporto 2007, volume X. Bologna: Ed. Centauro, 2007
7. Osservatorio ARNO-CINECA. Osservatorio ARNO Diabete - Analisi di dieci anni di prescrizioni. Rapporto 2007, Volume XI; disponibile su <http://osservatorioarno.cineca.org/convegni/diabete>