



Utilizzo epidemiologico di archivi sanitari elettronici correnti
Un'esperienza di standardizzazione delle procedure
per la stima di frequenza di alcune malattie in diverse aree italiane

Exploiting electronic health archives for epidemiological purposes
An experience using a standardized approach to estimate diseases in various areas of Italy

A cura di:

Lorenzo Simonato, Cristina Canova, Gianni Corrao, Giuseppe Costa, Roberta Tessari

**Composizione del Gruppo di lavoro nazionale AIE-SISMEC
su "Utilizzo epidemiologico di archivi sanitari elettronici di popolazione"**

Ileana Baldi (Torino), Daniela Balzi (Firenze), Alessandro Barchielli (Firenze), Giuseppe Battistella (Treviso), Cristina Canova (Padova), Giulia Cesaroni (Roma), Giovanni Corrao (Milano), Francesca Collini (Firenze), Susanna Conti (Roma), Giuseppe Costa (Torino), Moreno Demaria (Torino), Carla Fornari (Milano), Annunziata Faustini (Roma), Claudia Galassi (Torino), Roberto Gnani (Torino), Andrea Inio (Venezia), Fabiana Madotto (Milano), Enrica Migliore (Torino), Giada Minelli (Roma), Michele Pellizzari (Thiene), Mariangela Protti (Pisa), Anna Romanelli (Pisa), Antonio Russo (Milano), Mario Saugo (Thiene), Lorenzo Simonato (Padova), Valeria Tancioni (Roma), Roberta Tessari (Venezia), Paolo Trerotoli (Bari), Alice Vianello (Padova), Maria Angela Vigotti (Taranto)

Hanno inoltre contribuito ai lavori (ringraziamenti)

Gioacchino Bellone (Pisa), Marina Bonfanti (Pisa), Antonella Bruni (Taranto), Massimiliano Bugiani (Torino), Enrico Camerin (Thiene), Silvia Cascini (Roma), Giancarlo Cesana (Milano), Francesco Chini (Roma), Ludmila Karaghiosoff (Torino), Sante Minerba (Taranto), Pasquale Pepe (Pisa), Roberta Picariello (Torino), Pavilio Piccioni (Torino), Mauro Raciti (Pisa), Stefano Rigon (Thiene)

Indice Contents

| | | | |
|--|----------------|---|----------------|
| Prefazione / Foreword | pag. 3 | 6. Stima della prevalenza di broncopneumopatia cronico-ostruttiva basata su dati sanitari correnti mediante l'uso di un algoritmo comune in differenti aree italiane | pag. 46 |
| <i>Benedetto Terracini</i> | | Chronic obstructive pulmonary disease prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy (<i>Faustini A. et al</i>) | |
| 1. Obiettivi, strumenti e metodi per un utilizzo epidemiologico di archivi sanitari elettronici correnti in diverse aree italiane | pag. 5 | 7. Stima della prevalenza dell'asma bronchiale basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane | pag. 56 |
| Objectives, tools and methods for an epidemiological use of electronic health archives in various areas of Italy (<i>Simonato L. & gruppo di lavoro</i>) | | Asthma prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy (<i>Tessari R. et al</i>) | |
| 2. Stima della prevalenza del diabete basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane | pag. 15 | 8. Stima della prevalenza di malattie polmonari cronico-ostruttive basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane | pag. 66 |
| Diabetes prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy (<i>Gnavi R. et al</i>) | | Obstructive lung disease prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy (<i>Migliore E. et al</i>) | |
| 3. Stima della prevalenza della cardiopatia ischemica basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane | pag. 22 | Conclusioni comuni ai capitoli sulle patologie respiratorie | pag. 78 |
| Ischemic heart disease prevalence estimated using a standard algorithm based on electronic health data in various areas of Italy (<i>Balzi D. et al</i>) | | | |
| 4. Stima dell'incidenza dell'infarto miocardico acuto basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane | pag. 30 | 9. Procedure di record linkage in epidemiologia: uno studio multicentrico italiano | pag. 79 |
| Acute myocardial infarction incidence estimated using a standard algorithm based on electronic health data in different areas of Italy (<i>Barchielli A. et al</i>) | | Record-linkage procedures in epidemiology: an Italian multicentre study (<i>Fornari C. et al</i>) | |
| 5. Stima dell'incidenza di ictus acuto basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane | pag. 38 | 10. Conclusioni / Conclusions | pag. 89 |
| Acute stroke incidence estimated using a standard algorithm based on electronic health data in various areas of Italy (<i>Tancioni V. et al</i>) | | | |
| | | Appendice / Annex | pag. 93 |
| | | Ricerca e sviluppo di algoritmi: la definizione di alcune patologie neoplastiche | |
| | | Research and development of algorithms: definition of some neoplastic pathologies | |

Prefazione

Foreword

Intorno al 1840 William Farr è stato tra i primi a prevedere che in Gran Bretagna le statistiche di mortalità del Registrar general avrebbero fornito le informazioni che stanno alla base dell'epidemiologia: chi viene colpito da una malattia, quando, dove e perché (*who, when, where and why*).

Erano gli stessi anni in cui il dottor Rigoni Stern, dai registri dei morti a Verona, inferiva con grande saggezza sui fattori di rischio per i tumori femminili. Le previsioni di William Farr si sono avverate, ma né egli né Rigoni Stern potevano immaginare la molteplicità di statistiche utilizzabili per l'indagine sanitaria che si sarebbero rese disponibili nel successivo secolo e mezzo, e neppure gli sviluppi tecnologici che – salvo il rispetto per i principi della confidenzialità – avrebbero portato tali statistiche alla portata di miriadi di ricercatori.

Le statistiche correnti maggiormente utilizzate in sanità rimangono quelle di mortalità: anche in questo ambito la disponibilità di archivi automatizzati e i progressi degli strumenti per la ricerca epidemiologica sono stati travolgenti.

In Italia, sono trascorsi meno di quarant'anni da quando hanno cominciato a rendersi disponibili su base elettronica i primi dati di mortalità per provincia. Erano tempi in cui la carenza informativa delle statistiche correnti veniva compensata dall'ingegno di chi voleva «spremerle» per cavarne indicazioni utili in sanità pubblica. Il migliore esempio in questo senso è stato il primo *Atlante italiano di mortalità per tumore* pubblicato dalla Lega italiana contro i tumori nel 1975.

Il termine «archivi sanitari elettronici» che è stato scelto dagli autori di questo supplemento di *Epidemiologia & Prevenzione*, a rigore, include qualsiasi database elettronico con dati di potenziale interesse sanitario. In realtà, gli archivi usati nel lavoro corrispondono a statistiche correnti, in quanto previste da normative nazionali o regionali. Non è una questione soltanto semantica: pro-

prio per il fatto di essere previste da una normativa, le statistiche correnti consentono qualche forma di interazione tra ricercatori e autorità di politica sanitaria.

I curatori del supplemento sottolineano un punto importante: è ben vero che molte statistiche correnti oggi disponibili sono state create per scopi amministrativi, o comunque diversi da quelli sanitari, e peccano quindi in precisione diagnostica. La loro precisione può tuttavia diventare più soddisfacente se le si integra con archivi riguardanti l'ambito più strettamente diagnostico (come quelli dei servizi di anatomia patologica o delle prestazioni specialistiche).

E' importante conoscere le potenzialità di questi archivi e saggiare la possibilità – a proposito di alcune patologie rilevanti – di integrare archivi di natura diversa. E' un esercizio che merita senz'altro di essere fatto, ma è evidente che non è ancora stata completata la definizione e la valutazione, caso per caso, di un gold standard per valutare l'attendibilità delle osservazioni.

L'esperienza qui descritta non è quindi da considerare un prodotto finito, bensì un inizio di ragionamento del milieu epidemiologico italiano (incidentalmente, esistono esperienze straniere confrontabili con quella qui descritta?).

Il fine ultimo dell'epidemiologia è fornire indicazioni per interventi preventivi e, più in generale, di proteggere la salute delle persone. Saranno necessari alcuni anni per valutare l'entità del contributo che, in questo senso, può derivare dai progressi della tecnologia elettronica. Nel frattempo, a fianco dello sviluppo degli archivi elettronici, è importante fare emergere le ipotesi rilevanti, per la salute della gente, che si vuole verificare attraverso l'uso degli archivi stessi.

Benedetto Terracini
direttore scientifico
Epidemiologia & Prevenzione