

PREMESSA: Lo sviluppo delle tecniche di miniaturizzazione dell'informazione ha reso in pratica illimitata la quantità di informazioni archiviabili in ogni settore.

Le potenzialità informatiche sviluppatesi negli ultimi 25 anni hanno, infatti, reso possibile l'archiviazione di una enorme quantità di informazioni, a scopi prevalentemente amministrativi o economici, come rimborsi fra Aziende Sanitarie, calcolo dei DRG, e proprio per questo raccolte con una certa assiduità; l'importanza di questi archivi dal punto di vista epidemiologico risiede nella presenza di informazioni concernenti anche l'ambito più strettamente diagnostico, affiancate a quelle anagrafiche.

Rientrano in queste fonti gli archivi di mortalità, delle schede di dimissione ospedaliera (SDO), dei referti di anatomia patologica, gli archivi delle prescrizioni farmaceutiche, delle visite specialistiche, etc.

La potenzialità e l'innovazione rappresentate dall'impiego di questi archivi in epidemiologia scaturiscono proprio dal passaggio ad una loro gestione integrata: le informazioni raccolte attraverso un canale informativo possono così essere controllate sulla base dell'incrocio con gli altri archivi e la stessa portata informativa delle basi di dati si accresce notevolmente.

Lo scopo generale di questo progetto è dunque lo sfruttamento di questo prezioso e facilmente disponibile apparato informativo: molteplici sono, infatti, le prospettive di utilizzo dei dati raccolti con sistematicità se organizzati in sistemi integrati

04 FEB. 2008

TITOLO SVILUPPO DI SISTEMI DI SORVEGLIANZA EPIDEMIOLOGICA BASATI SULL'UTILIZZO DEGLI ARCHIVI ELETTRONICI SANITARI IN AMBITO DI SANITA' PUBBLICA: UNO STUDIO PILOTA IN ALCUNE ASL SELEZIONATE

ANALISI STRUTTURATA DEL PROBLEMA

Problema

Il governo della sanità richiede sempre più strumenti tempestivi ed affidabili in grado di valutare lo stato di salute della popolazione, ovvero della distribuzione delle malattie in essa.

In assenza di strumenti innovativi e moderni la promozione della salute ed il governo della sanità possono basarsi solo su dati di mortalità, su dati epidemiologici derivanti da indagini campionarie o da studi prospettici aventi come oggetto di indagine patologie specifiche limitando così lo spettro d'azione e l'efficacia degli interventi a livello della popolazione degli assistiti.

Grazie soprattutto agli enormi progressi fatti nel campo della miniaturizzazione dell'informazione si è resa possibile l'archiviazione elettronica di enormi quantità di informazioni nel campo della sanità. Sono ormai disponibili nella maggior parte del territorio nazionale estese basi dati all'interno dei presidi ospedalieri e delle istituzioni sanitarie in genere.

Gli archivi più frequentemente esistenti riguardano le schede di ricovero e di dimissione ospedaliera (SDO), le prestazioni ambulatoriali, i referti anatomo-patologici, i certificati di morte, le schede di consumo farmaci e, più recentemente, i dati dai laboratori e dalla diagnostica strumentale. L'uso di questa grande massa di dati è prevalentemente amministrativo, mentre ogni tentativo indirizzato ad un uso epidemiologico di questi dati è stato finora limitato. La potenzialità e l'innovazione rappresentate dall'impiego di questi archivi in epidemiologia scaturiscono proprio dal passaggio ad una loro gestione integrata: le informazioni raccolte attraverso un canale informativo possono così essere controllate sulla base dell'incrocio con gli altri archivi e la stessa portata informativa delle basi di dati si accresce notevolmente.

La sfida che si presenta ora in epidemiologia sanitaria è dunque lo sfruttamento di questo prezioso apparato informativo: molteplici sono, infatti, le prospettive di utilizzo dei dati raccolti con sistematicità se organizzati in sistemi integrati (1-5).

Spiegazioni plausibili

La scarsità di dati epidemiologici utilizzati in sanità pubblica è da imputare a più fattori:

- La carenza di cultura epidemiologica dei decisori in campo sanitario. Raramente infatti le figure istituzionali con responsabilità direttive in campo sanitario hanno una solida preparazione epidemiologica che di supporto per i processi valutativi e decisionali.
- Il ritardo con cui i dati epidemiologici esistenti (mortalità, registri tumori, indagini campionarie, etc) vengono spesso resi disponibili.
- La mancanza di metodi e strumenti consolidati che permettano un diffuso utilizzo a fini epidemiologici degli archivi sanitari amministrativi.

Soluzioni proposte

Nel periodo 2005-2007 è stato reso operativo grazie alla collaborazione fra l'Associazione Italiana di Epidemiologia (AIE) e la Società Italiana di Statistica Medica e Epidemiologia Clinica (SISMEC) un gruppo di lavoro composto da esperti nel campo dell'utilizzo di fonti elettroniche sanitarie che si è prefisso il compito di sviluppare dei comuni modelli di utilizzo di fonti correnti allo scopo di rendere confrontabili nelle territorio nazionale stime di incidenza e/o prevalenza di alcune delle maggiori patologie che colpiscono la popolazione generale.

Le patologie individuate sono state: cardiopatie ischemiche, infarto del miocardio, ictus, asma, broncopneumopatia cronica ostruttiva, malattia polmonare cronica ostruttiva, diabete. Sono stati elaborati per ciascuna di queste patologie specifici algoritmi in grado di definire i casi basandosi su diverse combinazioni delle seguenti fonti correnti: schede di dimissione ospedaliera, prescrizioni farmaci, certificati di morte, esenzioni ticket.

Per queste patologie sono stati utilizzati metodi comuni di trattamento dati, analisi e controllo di alcuni parametri di funzionamento degli algoritmi al fine di ridurre al massimo le differenze imputabili ai diversi approcci utilizzati nel produrre stime.

I risultati di questo lavoro, seppur con qualche disomogeneità, sono stati valutati positivamente e costituiscono un patrimonio di strumenti e conoscenze da sfruttare nel campo della sanità pubblica.

I metodi messi a disposizione rappresentano potenzialmente uno strumento relativamente semplice, e quindi a basso costo, per avere stime di incidenza e/o prevalenza non su base campionaria, ma sull'intera popolazione residente nelle diverse aree del territorio nazionale. L'unità periferica ideale per questo utilizzo sono le ASL o aggregazione di esse in quanto depositarie dei dati nominali necessari per le stime, e in quanto unità di base di funzionamento del SSN.

In questo ambito verrà verificata la possibilità di attivare sistemi epidemiologici integrati che si avvalgono dell'interazione fra fonti informative e di testare la applicabilità di trasferire in alcune ASL campione, selezionate sulla base della loro disponibilità di archivi elettronici di buona completezza e qualità, gli strumenti epidemiologici e informatici elaborati dal gruppo sotto la supervisione di un gruppo di esperti coordinati dal Dipartimento di Medicina Ambientale e Sanità Pubblica dell'Università di Padova.

Saranno inoltre supportati ulteriori sviluppi dell'attività di Ricerca e Sviluppo del gruppo collaborativo nella direzione sia dell'esplorazione di algoritmi per altre patologie, che dell'elaborazione di metodi per il controllo di qualità e la validazione della casistica.

Il progetto non va quindi considerato a termine in considerazione della necessità di adattare questi sistemi integrati al mutare quantitativo e qualitativo delle fonti informative.

Fattibilità

Il progetto si avvarrà dell'esperienza e delle professionalità statistiche ed epidemiologiche del gruppo nazionale AIE/SISMEC e dell'esperienza del Laboratorio per lo Sviluppo di Sistemi Epidemiologici Integrati dell'Università di Padova che ha attivato sistemi di sorveglianza in alcune ASL del Veneto. La scelta delle ASL, e conseguentemente delle Regioni di appartenenza, da includere nel progetto verrà decisa sulla base della valutazione del livello organizzativo degli archivi sanitari e della disponibilità e professionalità del personale da coinvolgere nelle ASL. L'ASL rappresenta l'unità operativa di base nella quale si svolge il progetto in parallelo con quella che è la sua funzione nel Servizio Sanitario Nazionale. Il progetto pilota si svolge nell'arco di due anni, ma l'attività di organizzazione sviluppo di sistemi di sorveglianza all'interno delle ASL, una volta a regime, dovrà ovviamente continuare al di là del periodo previsto.

Criticità

La qualità e la completezza dei dati archiviati dalle ASL possono essere fortemente deficitarie, oppure, anche in presenza di dati soddisfacenti, il periodo temporale della loro disponibilità può essere insufficiente per condurre analisi di adeguata potenza statistica determinando conseguentemente l'esclusione dal progetto.

Si dovranno per questi motivi prendere in considerazione un numero di ASL superiore a quello previsto del progetto privilegiando inoltre una distribuzione geografica che non penalizzi alcune aree, come il meridione, rispetto al resto del territorio nazionale.

Bibliografia

[1] 1° Rapporto sulla salute dei veneziani 1999-2003. <http://www.ulss12.ve.it/>

[2] Exploiting existing data for health research. International Conference. St. Andrews UK 18-20 September 2007

[2] Motheral BR, Fairman KA. The use of claims databases for outcome research: rationale, challenges and strategies. *Clinical Therapeutics* 1997; 19: 346-66.

[4] Osborne ML, Vollmer WM, Johnson RE, et al. Use of automated prescription database to identify individuals with asthma. *J Clin Epidemiol* 1995; 48: 1393-7.

[5] Strom BL, Carson JL. Use of automated databases for pharmacoepidemiology research. *Epidemiol Rev* 1990; 12:87-107.

Allegato 2

OBIETTIVO GENERALE

Attivare un sistema di sorveglianza epidemiologica basato su meccanismi di integrazione di dati correnti sanitari

OBIETTIVO SPECIFICO 1

Valutare la fattibilità di attivazione del sistema di sorveglianza proposto in alcune ASL selezionate esaminando gli archivi sanitari informatizzati necessari allo scopo sotto il profilo della disponibilità temporale, la completezza e la qualità

OBIETTIVO SPECIFICO 2

Trasferire al personale delle ASL coinvolte gli strumenti metodologici ed informatici necessari all'integrazione delle fonti correnti

OBIETTIVO SPECIFICO 3

Attivazione del sistema di sorveglianza e produzione dei risultati epidemiologici per le maggiori patologie presenti nella popolazione studiata

OBIETTIVO SPECIFICO 4

Coordinamento di un gruppo di esperti con compiti di ricerca e sviluppo dei metodi

PIANO DI VALUTAZIONE PER OGNI OBIETTIVO SPECIFICO

Obiettivo generale	Attivare un sistema di sorveglianza epidemiologica basato su meccanismi di integrazione di dati correnti sanitari	
Obiettivo specifico 1	Valutare la fattibilità di attivazione del sistema di sorveglianza proposto in alcune ASL selezionate esaminando gli archivi sanitari informatizzati necessari allo scopo sotto il profilo della disponibilità temporale, la completezza e la qualità. Architetture il data-base del sistema di sorveglianza epidemiologica previa valutazione positiva.	
Indicatore di risultato	Relazione sulla situazione dei singoli archivi sanitari e valutazione della loro utilizzabilità per il sistema di sorveglianza epidemiologica	
Standard di risultato	Statistiche descrittive delle distribuzioni delle variabili esaminate e risultati dei controlli logici di qualità effettuati	
Azione	Indicatore/i di processo	Standard di processo
Analisi dei dati contenuti negli archivi sanitari	Estrazione dei dati dalle fonti correnti	100%
	Valutazione della completezza e della qualità	100%
	Preparazione di una relazione specifica per ogni ASL coinvolta	1

Obiettivo generale	Attivare un sistema di sorveglianza epidemiologica basato su meccanismi di integrazione di dati correnti sanitari	
Obiettivo specifico 2	Trasferire al personale delle ASL coinvolto gli strumenti metodologici ed informatici necessari all'integrazione delle fonti correnti	
Indicatore di risultato	Nuove conoscenze acquisite dal personale coinvolto attraverso corsi e/o addestramento pratico personalizzato	
Standard di risultato	Autonomia operativa del personale ASL nella gestione del sistema di sorveglianza	
Azione	Indicatore/i di processo	Standard di processo
Addestramento del personale ASL	Organizzazione corsi e seminari nelle ASL	1 per ogni ASL
	Verifica pratica delle capacità acquisite	100%

Obiettivo generale	Attivare un sistema di sorveglianza epidemiologica basato su meccanismi di integrazione di dati correnti sanitari	
Obiettivo specifico 3	Attivazione del sistema di sorveglianza e produzione dei risultati epidemiologici per le maggiori patologie presenti nella popolazione studiata	
Indicatore di risultato	Realizzazione del data-base necessario per l'attivazione del sistema di sorveglianza	
Standard di risultato	Disponibilità di stime di frequenza delle maggiori patologie nella popolazione	
Azione	Indicatore/i di processo	Standard di processo
Analisi descrittive dei dati integrati	Caratteristiche epidemiologiche della popolazione degli assistiti	100%
	Stime di frequenza delle principali malattie	90%

Obiettivo generale	Attivare un sistema di sorveglianza epidemiologica basato su meccanismi di integrazione di dati correnti sanitari	
Obiettivo specifico 4	Coordinamento di un gruppo di esperti con compiti di ricerca e sviluppo dei metodi	
Indicatore di risultato	Aggiornamento di algoritmi esistenti e/o proposta di algoritmi di definizione di nuove patologie	
Standard di risultato	Stime di frequenza aggiornate o nuove precedentemente non disponibili	
Azione	Indicatore/i di processo	Standard di processo
Produzione di stime aggiornate o nuove	Disponibilità di un rapporto di sintesi dei risultati ottenuti	1

PIANO FINANZIARIO GENERALE

Risorse	1° anno *	2° anno *	totale
Personale	40,000	20,000	60,000
Beni e servizi	2,500	2,500	5,000
Missioni	12,500	12,500	25,000
Spese generali	5,000	5,000	10,000
Totale	60,000	40,000	100,000

* a seconda del progetto (annuale, biennale, triennale) le risorse finanziarie dovranno essere divise per singolo anno

INTEGRAZIONE DEL PROGETTO ESECUTIVO: SISTEMI DI SORVEGLIANZA EPIDEMIOLOGICA BASATI SULL'UTILIZZO DEGLI ARCHIVI ELETTRONICI SANITARI IN AMBITO DI SANITA' PUBBLICA: UNO STUDIO PILOTA IN ALCUNE ASL SELEZIONATE

Obiettivo 1 : Gli archivi sanitari che verranno valutati sono quelli dei certificati di morte informatizzati e codificati secondo codice ICD, delle Schede di Dimissione Ospedaliera con informazioni diagnostiche multiple informatizzate e codificate secondo codice ICD, delle schede della Mobilità Passiva riguardanti le prestazioni agli assistiti fornite da altre ASL, delle diagnosi dei referti cito-istopatologici informatizzate e codificate secondo codice ICD-O o SNOMED (per questi codici verranno forniti programmi di transcodifica), delle prescrizioni farmaceutiche informatizzate e codificate secondo codice ATC, delle esenzioni ticket informatizzate e codificate. Altre fonti esistenti come l'archivio della specialistica o l'archivio delle pensioni di invalidità non sono attualmente informatizzate e codificate, ma possono costituire un prezioso patrimonio per ulteriori integrazioni e per le verifiche di qualità, per cui verranno anch'esse esaminate e valutate.

Infine verrà verificata la completezza ed il livello qualitativo dell'anagrafe sanitaria che è elemento indispensabile per l'attivazione del sistema in quanto costituisce la matrice di aggancio di tutte le informazioni raccolte. Tutte le informazioni sanitarie dagli archivi menzionati dovranno infatti essere incrociate con l'anagrafe sanitaria attraverso operazioni di record-linkage deterministico per determinare l'effettiva appartenenza dei soggetti con prestazioni all'ASL in studio.

Per tutte le fonti sopramenzionate verranno condotti un'analisi della completezza e altri controlli logici di qualità per poter selezionare le ASL presso le quali condurre lo studio pilota.

Considerando le limitate risorse disponibili si pensa di poter intervenire in 3-4 ASL da selezionare da un gruppo di una decina di eligibili.

Criteri di selezione saranno, oltre ovviamente all'interesse dei Direttori Generali, la qualità e la completezza dei dati, l'intervallo temporale entro il quale sono disponibili, e la disponibilità di personale da addestrare.

Verranno escluse le aree già incluse nell'attività del gruppo di lavoro AIE/SISMEC o in altri progetti simili, mentre verrà tenuto in considerazione il rapporto nord-sud nella collocazione geografica.

Verranno inoltre privilegiate le ASL nelle quali sia possibile confrontare i risultati derivanti dal progetto con quelli da fonti o sistemi di raccolta dati indipendenti e metodologicamente diversi allo scopo di validare le stime prodotte.

Obiettivo 2 : Il personale delle ASL coinvolto nel progetto opera generalmente all'interno dei CED (centri elaborazione dati) o SIL (sistemi informativi locali) che gestiscono gli archivi elettronici con competenze statistiche e/o informatiche. Non raramente si tratta di personale con background diverso che ha acquisito competenze sul campo. E' opportuno che all'interno delle ASL vengano coinvolte anche le figure con responsabilità epidemiologiche all'interno dei Dipartimenti Prevenzione o in Unità di Epidemiologia qualora esistano.

Per il personale interessato nelle ASL selezionate verrà condotto un addestramento su due livelli:

- a) uno generale avente lo scopo di far acquisire al personale le basi metodologiche del progetto, i livelli organizzativi necessari, le caratteristiche dei dati e le procedure per standardizzarne la gestione. Questo livello verrà organizzato centralmente per tutte le ASL coinvolte e verrà integrato da esperienze pratiche di utilizzo degli strumenti informatici standard;
- b) uno specifico che verrà invece svolto presso le ASL e mirante ad affrontare nel particolare gli aspetti legati alla specificità dei sistemi operativi che sono contraddistinti da notevole variabilità da ASL a ASL.

Obiettivo 3 : la prima fase di attivazione del sistema di sorveglianza consiste nell'aggancio attraverso record-linkage delle schede elettroniche nelle quali è condensata l'informazione sanitaria agli assistibili usando come chiavi di aggancio il cognome e nome, il codice sanitario, e la data di nascita.

Un processo step-wise di incroci di informazioni sarà quindi sviluppato utilizzando le fonti informative computerizzate e gli archivi di popolazione che rimangono la matrice di qualunque sistema di incrocio. Questa fase costituisce la **validazione** dell'appartenenza delle informazioni agli individui assistiti dell'ASL. Molto utile sarà in questa fase poter accedere alle anagrafe comunali per poter effettuare delle verifiche.

A questa fase seguirà la fase dell'**allineamento** che consiste nell'assegnare tutte le informazioni elettroniche, ovvero tutti gli eventi sanitari disponibili, provenienti dalle varie fonti ad ogni singolo soggetto ricostruendo così la storia individuale degli eventi patologici.

La terza fase del processo è il **riconoscimento** delle malattie attraverso l'elaborazione di algoritmi patologia-specifici che trasferiscono a livello informatico le conoscenze e gli algoritmi decisionale che ispirano qualunque diagnosi di malattia.

Gli algoritmi verranno testati in differenti situazioni nosologiche e nelle varie ASL per valutare la consistenza dei risultati. Programmi informatici standard appositamente disegnati saranno messi a disposizione dei centri collaborativi per testarne la consistenza e valutare la variabilità dei risultati in rapporto alla stabilità delle stime ottenute, all'analisi delle diverse combinazioni delle fonti ed al loro apporto specifico ed assoluto.

Le caratteristiche delle stime così ottenute per le patologie più importanti verranno analizzate sotto il profilo epidemiologico e criticamente valutate in una prospettiva di uso da parte delle pubbliche amministrazioni.

Si assume che procedure intelligenti orientate alla registrazione in automatico delle malattie siano caratterizzate da una maggiore sensibilità e da una minore specificità a confronto con i sistemi di registrazione tradizionali. Diviene quindi cruciale mettere in opera parallelamente al processo in automatico procedure di controllo di qualità incaricate di definire e dimensionare la proporzione dell'errore prodotto da questi sistemi, che dovrebbe ricadere soprattutto nell'ambito dei falsi positivi. Questi controlli di qualità dovrebbero essere regolarmente applicati attraverso appropriate tecniche di campionamento allo scopo di migliorare gli algoritmi di risoluzione e misurare con precisione l'errore prodotto dal sistema automatico.

Basi dati computerizzate in campo medico, ove disponibili verranno valutate per la completezza e la qualità dell'informazione esistente.

Obiettivo 4 : In parallelo con le attività dei precedenti obiettivi specifici verrà costituito un gruppo di esperti invitati a partecipare sulla base documentata degli studi condotti specificatamente in questo campo.

Il gruppo così costituito avrà tre compiti principali: valutare criticamente i progressi ed i risultati delle attività descritte negli obiettivi 1-3; elaborare algoritmi su patologie finora non considerate e/o proporre modifiche degli algoritmi esistenti; elaborare strategie di validazione delle stime prodotte attraverso gli algoritmi