



UNIVERSITA' DEGLI STUDI DI PADOVA
FACOLTA' DI SCIENZE STATISTICHE

CORSO DI LAUREA TRIENNALE IN
STATISTICA, POPOLAZIONE E SOCIETA'

Tesi di Laurea

VALUTAZIONE CRITICA DI UN SISTEMA EPIDEMIOLOGICO
INTEGRATO

Potenzialità e limiti dell'utilizzo integrato di archivi elettronici di
popolazione per il monitoraggio della frequenza delle malattie

Relatore: CH.MO Prof. Simonato Lorenzo

Laureanda: Vianello Alice
Matricola: 518690-SPT

ANNO ACCADEMICO 2006-2007

INDICE

Introduzione	pag. 5
Materiali e metodi: il SEI di Venezia	pag. 9
Risultati: il gruppo di lavoro nazionale AIE-SISMEC	pag. 23
Discussione	pag. 89
Conclusioni	pag. 99
Bibliografia	pag. 101

Capitolo 1

INTRODUZIONE

Il momento attuale si caratterizza in ambiti diversi per un interesse crescente rivolto alla programmazione: questa attività si prefigge di impiegare le risorse a disposizione con oculatezza, fornendo nei tempi ritenuti opportuni le risposte più adeguate ai problemi che si presentano nel presente.

Programmare implica partire da una lettura approfondita della situazione attuale per predisporre in modo ordinato le azioni future che, a breve o lungo termine, andranno a modificarla.

Fondamento imprescindibile è dunque una panoramica della realtà presente, che ne metta in luce le dinamiche in atto, evidenziandone gli snodi più problematici sui quali si rende necessario un intervento.

In ambito sanitario, la programmazione riveste un'importanza cruciale per il futuro sviluppo del sistema: le risorse notoriamente limitate e i bisogni della popolazione tendenzialmente crescenti impongono di affrontare le problematiche in modo globale, con raziocinio e competenza. Soprattutto in quest'ambito, infatti, gli operatori del settore sono tenuti a pianificare gli interventi sulla base dei bisogni, essendo la salute una sfera che incontra particolare sensibilità nella popolazione ed imponendo le disponibilità finanziarie di ridurre al minimo le inefficienze, per rispondere al meglio alla domanda sanitaria.

Per poter effettuare una programmazione è necessario che sia nota la situazione di partenza, affinché gli interventi siano mirati, efficaci ed efficienti. Nel contesto sanitario, l'informazione essenziale dalla quale partire è la proporzione di malati sulla popolazione e la tipologia delle malattie che interessano gli assistiti.

Per intervenire sullo stato di salute della popolazione bisogna, infatti, prevalentemente combattere le malattie che l'affliggono, pianificando attività di cura, prevenzione e promozione.

In questo ambito si inserisce l'epidemiologia descrittiva, che è in grado di fornire le informazioni richieste al fine di analizzare la situazione sulla quale si intende intervenire.

Tasso di incidenza e di prevalenza delle malattie sulla popolazione degli assistiti sono, dunque, le due informazioni essenziali che l'epidemiologia deve mettere a disposizione come strumenti di partenza per il successivo lavoro di programmazione sanitaria.

L'incidenza misura la proporzione di nuovi malati in una popolazione in un intervallo temporale definito $(t, t+1)$ ¹.

La prevalenza fornisce la proporzione di malati in una popolazione in un preciso momento (t) ¹.

Come pervenire alle informazioni su incidenza e prevalenza di malattie in una popolazione? Questa è essenzialmente la domanda fondamentale attorno alla quale ruota questo lavoro.

Attualmente vi sono diversi metodi che permettono di stimare queste due grandezze: si possono ottenere i dati rivolgendosi direttamente agli individui, chiedendo loro da quali malattie sono afflitti; in questo caso la rilevazione potrà essere censuaria, interessando tutti gli assistiti, oppure campionaria, selezionando un campione rappresentativo della popolazione. Questo secondo sistema è impiegato dall'ISTAT nell'Indagine sulla salute nell'ambito dell'Indagine Multiscopo sulle famiglie². I principali limiti di questo metodo di raccolta sono l'attendibilità delle informazioni, poiché trattandosi di una tematica estremamente sensibile i diretti interessati spesso oppongono resistenza alla richiesta di informazioni, omettendole o

¹ Cfr. Bibliografia: Simonato L., Centro Oncologico Regionale

² Cfr. Bibliografia: ISTAT 2002; ISTAT, 2005

falsificandole, e i tempi di raccolta ed elaborazione delle informazioni, decisamente lunghi, specialmente nel caso di raccolta censuaria.

Si possono altrimenti costituire appositi registri di patologia, che raccolgono i dati relativi ad ogni nuovo caso della specifica malattia nella popolazione. Le informazioni ricavabili da questo sistema di raccolta sono solitamente molto dettagliate e attendibili, poiché la fonte del dato non è più la persona, bensì lo specialista che effettua la diagnosi. I registri sono però sistemi complessi, che richiedono molto lavoro di gestione e si limitano solitamente ad una patologia o ad un gruppo di patologie (ad esempio, Registro Diabetici e Registro Tumori) e finora sono stati estesi solamente ad alcune malattie.

Le fonti di dati sanitari non si esauriscono però certamente in malati e medici curanti: 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.

L'unica registrazione resa obbligatoria a livello nazionale è quella dei certificati di morte. La prassi di archiviazione di dati sanitari è comunque diffusa, almeno in parte, presso tutte le Aziende Sanitarie, e sottoposta a specifici regolamenti regionali; gli archivi sono generalmente a se stanti, organizzati e gestiti separatamente.

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 quanto, potenzialmente, facilmente disponibile apparato informativo: molteplici sono, infatti, le prospettive di utilizzo dei dati raccolti con sistematicità se organizzati in sistemi integrati.

Nei capitoli successivi si affronteranno la presentazione dettagliata delle caratteristiche di un sistema epidemiologico integrato (capitolo 2), la discussione dei risultati ottenuti da un gruppo di lavoro nazionale sull'utilizzo di archivi elettronici per la stima di incidenza e prevalenza di alcune malattie sulla popolazione (capitolo 3), per arrivare infine all'aspetto più interessante di questo lavoro, un contributo di discussione critica del sistema nella prospettiva di sviluppo futuro (capitolo 4).

Capitolo 2

MATERIALI E METODI: IL *SEI* DI VENEZIA

Negli ultimi anni si sono sviluppate parallelamente diverse dinamiche nell'ambito della emergente programmazione sanitaria: una crescente richiesta di informazioni epidemiologiche per poter pianificare servizi in risposta a bisogni della popolazione reali e non semplicemente supposti; una progressiva diminuzione delle risorse economiche delle Aziende Sanitarie, che impone estremo rigore nella spesa; una crescente disponibilità di informazioni codificate e memorizzate anche nell'ambito sanitario, che ha permesso la nascita diffusa di archivi di dati.

La presenza di questi fattori pone al momento attuale una duplice prospettiva: considerare le diverse tendenze in modo a sé stante, prospettandone uno sviluppo parallelo e dunque separato o alternativamente integrare queste tendenze, per perseguire congiuntamente l'obiettivo comune: il miglior funzionamento del sistema sanitario che si possa ottenere partendo dalle risorse e dalle problematiche attuali per rispondere adeguatamente alle esigenze di salute della popolazione.

La strada che appare come più complessa ma con una prospettiva più lungimirante è sicuramente la seconda: si rende necessaria una sintesi fra le dinamiche emerse, nell'interesse dell'intero sistema sanitario e per rispondere in modo realmente adeguato ai bisogni della popolazione da esso assistita.

Come è possibile utilizzare le dinamiche in atto in sinergia, per far emergere i vantaggi di ognuna e limitarne contemporaneamente gli svantaggi?

Verrà di seguito presentata una struttura innovativa, pensata e realizzata nell'Azienda ULSS veneziana, proprio per rispondere a queste esigenze: il Sistema Epidemiologico Integrato (SEI).

Gli archivi elettronici

L'Azienda ULSS veneziana, che copre le popolazioni di Venezia, Mestre e Comuni minori limitrofi (Marcon, Quarto d'Altino e Cavallino), dispone, come la maggior parte delle Aziende Sanitarie, di molteplici archivi elettronici per la raccolta di dati relativi agli assistiti; la finalità principale per cui sono stati introdotti questi archivi è di carattere amministrativo, in quanto si rendeva necessario un agile strumento per il calcolo delle spese sostenute sulla base dei servizi erogati, per le richieste di rimborso, per gestire pagamenti e recuperi.

Esiste un archivio che registra tutti i decessi avvenuti sul territorio dell'Azienda, sia riguardanti gli assistiti dell'AULSS che esterni; ad ogni ricovero, ordinario o in regime di day-hospital, segue l'archiviazione di una scheda di dimissione ospedaliera (SDO) contenente informazioni riguardanti le diagnosi di ammissione e dimissione, principali ed eventualmente secondarie, gli spostamenti tra Reparti ospedalieri e le prestazioni erogate durante la permanenza presso la struttura; i dati relativi ai ricoveri di assistiti dell'Azienda avvenuti presso strutture afferenti ad altre Aziende Sanitarie sono forniti da queste ultime e registrati nell'archivio denominato "mobilità passiva"; i referti di anatomia patologica costituiscono una base dati separata, contenente i dati relativi alle diagnosi istologiche, citologiche ed autoptiche sui tessuti in seguito ad interventi chirurgici, in particolar modo quelli per contrastare le neoplasie; le prescrizioni farmaceutiche sono materialmente raccolte dalle farmacie e inviate poi all'ufficio dell'Azienda Sanitaria preposto alla loro registrazione; di recente anche le visite specialistiche vengono registrate; esiste infine un archivio delle esenzioni ticket, alle quali hanno diritto malati di patologie croniche il cui trattamento richiede il ricorso frequente a farmaci essenziali molto costosi.

I dati così raccolti si riferiscono, nominalmente o attraverso un codice personale, all'assistito che ha usufruito del servizio sanitario e le variabili registrate sono piuttosto numerose, di carattere sia anagrafico che diagnostico-sanitario.

Concretamente, ogni record in questi archivi è riferito al paziente al quale è stato erogato il servizio e l'informazione relativa alla tipologia di assistenza è codificata secondo classificazioni specifiche. Così, la causa del decesso nei certificati di morte e la diagnosi all'ammissione e alla dimissione nelle SDO sono codificate secondo la Classificazione internazionale delle malattie, traumatismi e cause di morte (ICD), redatta e periodicamente rivista dall'Organizzazione mondiale della Sanità³; i farmaci vengono registrati con codifica ATC; i referti di anatomia patologica sono codificati secondo il criterio SNOMED⁴.

Infine, l'archivio che contiene tutte le informazioni anagrafiche, nonché il codice personale, degli assistiti dell'Azienda è l'Anagrafe Sanitaria, elenco aggiornato degli individui che costituiscono la popolazione alla quale sono rivolti i servizi erogati dall'AULSS.

La ricchezza delle informazioni contenute nei diversi archivi e la riferibilità al singolo assistito dei dati raccolti sono le caratteristiche essenziali sulle quali si basa il Sistema Epidemiologico Integrato (SEI).

Il presupposto fondamentale è che i dati disponibili si possano utilizzare per scopi diversi da quelli per i quali vengono raccolti: gli archivi amministrativi si presentano in questa luce come preziose e irrinunciabili fonti informative alle quali l'epidemiologia può attingere.

³ Cfr. Bibliografia: ISTAT, Revisione 1975. Sta ora avvenendo il passaggio dalla IX alla X versione della Classificazione ICD.

⁴ *Systemised Nomenclature of Medicine*

Caratteristiche generali del SEI

Il SEI è un sistema di raccordo fra i diversi archivi, sopra elencati, contenenti dati raccolti presso l'Azienda ULSS veneziana.

I diversi archivi sono stati attivati in tempi differenti: i primi dati dei referti di anatomia patologica risalgono al 1981, mentre l'archivio delle SDO al 1982; l'archivio dei certificati di morte (CM) e l'Anagrafe Sanitaria contengono i dati completi per tutti gli assistiti dal 1994; le prescrizioni farmaceutiche sono disponibili dal 1998.

La banca dati del SEI è stata costituita nel 2002 per rispondere all'esigenza di supporto empirico alla programmazione sanitaria e per sfruttare la disponibilità informativa per finalità epidemiologiche; essa assolve diverse funzioni:

- descrizione della struttura e delle dinamiche che interessano la popolazione degli assistiti e le malattie nel tempo;
- supporto all'attività decisionale dell'Azienda nell'ambito della sanità pubblica;
- sorveglianza delle dinamiche correlate a fattori ambientali e professionali, attraverso il confronto e l'interazione con dati forniti da altri Enti (Comune, ARPAV, ecc.);
- attività di epidemiologia analitica attraverso l'individuazione della casistica per approntare studi eziologici di follow-up.

Questa struttura permette di interrogare la banca dati integrata ricavando informazioni raccolte da fonti diverse, che riferendosi agli stessi pazienti permettono di ottenere un quadro più articolato delle situazioni di ricorso al sistema sanitario, accrescendo la portata informativa posseduta singolarmente dagli archivi.

L'intera popolazione veneziana costituisce così una coorte che può essere seguita nel tempo e sulla quale si possono condurre molteplici studi epidemiologici, senza dover progettare indagini campionarie ad hoc per ricavare informazioni.

Il SEI può dunque essere qualificato come un sistema epidemiologico di sorveglianza sulla popolazione veneziana; l'attività di monitoraggio non è altresì limitata ad un periodo temporale, ma grazie alla sistematicità dell'integrazione degli archivi può essere agevolmente condotta senza soluzione di continuità, poiché l'aggiornamento di ogni singolo archivio comporta automaticamente l'aggiornamento dell'intera banca dati.

Attività compiute utilizzando il SEI

Mediante il SEI viene svolta un'attività ordinaria di produzione di dati descrittivi sugli assistiti dell'AULSS 12: nel 2005 è stato pubblicato il I° Rapporto sullo stato di salute della popolazione veneziana⁵ con dati del periodo 1998-2003, contenente più di 4000 stime di incidenza di una selezione di 50 patologie definite secondo la classificazione ICD-IX, distinte per i due sessi, i 5 anni di stima e le 2 aree geografiche, per confrontare il centro storico di Venezia con Mestre, zona di terraferma.

Essendo possibile ottenere il dato relativo all'indirizzo di residenza di ogni assistito dagli archivi comunali, utilizzando il software GIS⁶ è inoltre possibile la georeferenziazione della banca dati: in questo modo si possono condurre interessanti studi ambientali e realizzare mappe del territorio come supporto ad indagini eziologiche localizzando i casi della malattia di interesse.

⁵ Cfr. Bibliografia: Lafisca S. et al., 2005

⁶ Sistemi Informativi Territoriali

Nel 2006 sono state realizzate mappe territoriali, presentate in una tesi di laurea⁷, della diffusione delle principali malattie per i 17 sestieri e quartieri in cui è suddivisa la popolazione dell'AULSS; è stato calcolato il rapporto standardizzato di incidenza (RSI) per 29 patologie nel periodo 1998-2003, per sesso e congiuntamente, assegnando colori diversi a seconda del valore assunto nell'area.

Negli scorsi anni sono state esplorate inoltre le potenzialità del sistema come strumento per effettuare studi epidemiologici analitici: è stato condotto uno studio indagando il rapporto fra asma nei bambini ed inquinamento⁸, considerando asmatici i soggetti in età inferiore a 14 anni ai quali fosse stato prescritto almeno una volta un farmaco broncodilatatore, e uno studio sulla relazione fra inquinamento da diossine e insorgenza di sarcomi dei tessuti molli nella popolazione di Venezia – Mestre⁹, sempre utilizzando fonti informatizzate.

Questi e altri studi sono stati attuabili grazie allo sfruttamento del SEI, che è stato uno strumento fondamentale, a seconda dello studio, nella definizione della casistica, nel conteggio dei malati e nella georeferenziazione dei casi.

L'Azienda ULSS di Venezia partecipa altresì ad un gruppo di lavoro nazionale costituitosi nel 2005 allo scopo di esplorare le potenzialità degli archivi elettronici di popolazione nella stima di incidenza e prevalenza delle malattie.

I metodi impiegati da questo gruppo di lavoro e i risultati ottenuti verranno illustrati nel capitolo 3.

⁷ Cfr. Bibliografia: Recaldin L., 2006

⁸ Cfr. Bibliografia: Regione Veneto, 2005

⁹ Cfr. Bibliografia: Tessari R., Canova C., Canal F. et al., 2006

Il record linkage

La fonte tendenzialmente di miglior qualità contenente i dati anagrafici dei cittadini residenti in un Comune è l'Anagrafe Comunale, con la quale si raffronta il corrispettivo archivio sanitario per un controllo sui propri dati; nel caso di Venezia, le due basi di dati non sono completamente sovrapponibili, poiché la popolazione dell'AULSS 12 è suddivisa su più comuni (cfr. pag. 10).

Fra gli archivi elettronici presentati in precedenza, l'Anagrafe Sanitaria viene considerata la fonte di riferimento (gold standard) per conoscere la popolazione degli assistibili e ne costituisce il tramite: i pazienti sono individuati da un codice personale di tessera sanitaria, riportato in ogni banca dati, che permette di implementare una procedura di record linkage fra le varie fonti di interesse.

Combinazioni con altre informazioni anagrafiche (nome, cognome e data di nascita) sono utilizzate come chiavi nell'estrazione dei record personali dai diversi archivi incrociati, per controllare possibili errori nei vari campi.

Per la selezione dei record nel SEI vengono impiegate in sequenza le seguenti chiavi (procedura *stepwise*):

- K1 tessera sanitaria, cognome, nome, data di nascita
- K2 tessera sanitaria, cognome, data di nascita
- K3 tessera sanitaria, nome, data di nascita
- K4 tessera sanitaria, cognome, nome
- K5 cognome, nome, data di nascita

La sequenza prevede il passaggio da una chiave più completa, che garantisce un elevato livello di sicurezza relativo alla corrispondenza effettiva fra record che vengono attribuiti alla stessa persona, a chiavi progressivamente meno complete e quindi più suscettibili di errori.

Le procedure di record linkage esistenti in letteratura sono distinte in deterministiche, semi-deterministiche e probabilistiche.

Nella procedura deterministica si considerano relativi allo stesso soggetto due record che risultano identici dal confronto fra campi identificativi riferiti alla stessa variabile. Gli errori di trascrizione dei dati negli archivi sono molto diffusi ed eterogenei, e questo comporta il non appaiamento di una quantità difficilmente stimabile di record riferiti alla stessa persona.

La procedura semi-deterministica si differenzia dalla precedente sul piano operativo, mantenendone però lo stesso impianto teorico: il confronto passa dall'intero campo identificativo ad un suo sottoinsieme, come ad esempio le prime 11 cifre del codice fiscale. Questa seconda procedura si basa sull'assunto che la probabilità di riscontrare errori aumenti all'aumentare della lunghezza del campo, e che questi siano maggiormente probabili nella registrazione delle parti finali del campo. Il criterio di scelta della parte da confrontare è però a cura dell'operatore o di chi progetta il sistema, e anche in questo caso il calcolo dell'errore è problematico.

Infine, la procedura probabilistica confronta ogni campo all'interno delle due chiavi stabilite e assegna un punteggio ad ogni confronto, positivo in caso di concordanza e negativo in caso di discordanza. Al punteggio è associato un peso sulla base dell'attendibilità del campo, definita come probabilità di commettere errori di registrazione nel campo considerato, e del suo potere discriminante, definito come la probabilità che la variabile assuma una stessa modalità. Ad esempio, essendo riscontrabili facilmente errori nella registrazione del nome di un soggetto e meno comunemente nell'anno di nascita, la concordanza fra i due campi contenenti il nome avrà un peso minore rispetto alla concordanza assegnata ai due campi relativi all'anno di nascita; il potere discriminante del codice fiscale, codice unico per ogni persona, è sicuramente maggiore rispetto a quello del cognome.

Il vantaggio più apprezzato della procedura probabilistica è la misura dell'entità dell'errore di linkage; per poter applicare questo metodo si rende però necessaria una potenza di calcolo non sempre disponibile e l'integrazione di più di due archivi ne rende ancor più problematica la fattibilità.

Nel SEI di Venezia si applica una procedura deterministica.

E' stato effettuato un controllo di qualità per valutare l'entità dell'errore di linkage tra l'Anagrafe Sanitaria e l'archivio delle SDO. Per tener conto anche della mobilità attiva¹⁰, si è incrociato l'archivio con l'Anagrafe Regionale.

La percentuale di record così appaiati è del 94%; tra il 6% di record non appaiati è stato estratto un campione casuale costituito da 260 casi (3% dei non appaiati), sul quale è stato condotto un controllo manuale.

Il 4,2% del campione è stato recuperato dall'incrocio con l'Anagrafe Sanitaria, il 2,3% dall'Anagrafe Regionale e il 10,4% è stato individuato tramite l'archivio degli Stranieri Temporaneamente Presenti (STP); dal controllo manuale, il rimanente 83% di records non è risultato rintracciabile negli archivi a disposizione: ci si attende comunque che una quota di persone che usufruiscono dei servizi sanitari non risieda nel territorio coperto dall'AULSS o nella Regione, essendo Venezia molto frequentata da turisti residenti fuori Regione o all'estero.

Dal controllo non emergono dunque particolari preoccupazioni sulla qualità del record linkage fra gli archivi di interesse; errori più frequenti sono stati riscontrati nei campi del nome e del cognome.

¹⁰ Prende il nome di "mobilità attiva" il flusso di ricoverati presso strutture dell'AULSS veneziana ma assistiti da altre ULSS.

Potenzialità e criticità del SEI

L'esperienza del Sistema Epidemiologico Integrato è finora limitata all'Azienda ULSS veneziana, necessariamente quindi le scelte metodologiche che sono state fatte non hanno potuto essere supportate dal confronto con altre esperienze a livello nazionale o internazionale.

Sia per questo motivo, sia per la presenza di criticità, il sistema deve essere considerato sperimentale e quindi largamente modificabile e perfezionabile.

Nel periodo di utilizzo del sistema si sono evidenziati rilevanti aspetti positivi del suo funzionamento.

L'organizzazione integrata degli archivi elettronici permette una invidiabile tempestività nel produrre i risultati, aspetto importante e spesso problematico nell'attività epidemiologica e programmatoria. La vicinanza temporale tra fenomeno studiato e pubblicazione dei risultati dell'analisi è infatti spesso cruciale affinché lo studio effettuato possa essere concretamente applicato, apportando un contributo incisivo all'interpretazione delle dinamiche studiate.

Caratteristica da non sottovalutare nell'ambito della ricerca scientifica è la disponibilità di un'enorme quantità di dati a costi estremamente esigui, trattandosi di informazioni archiviate di routine, che ridurrebbe la necessità di progettare indagini ad hoc; i dati sono, inoltre, costantemente disponibili nel tempo sotto forma di flusso istituzionalizzato, con le garanzie che queste due caratteristiche presentano: confrontabilità spazio-temporale, riproducibilità delle analisi, regolarità e utilizzo di procedure standard nella raccolta dei dati.

Gli algoritmi per l'integrazione degli archivi e la definizione della casistica sono strutturati in forma relativamente semplice e, una volta collaudati, applicabili agevolmente e tempestivamente ai dati.

Infine, l'integrazione di diverse fonti permette un controllo trasversale dei dati contenuti nei singoli archivi, ad esempio attraverso controlli logici interni: dall'incrocio delle informazioni può infatti emergere un quadro coerente della situazione della malattia dell'assistito, oppure possono evidenziarsi difformità che dovranno portare alla ricerca dell'errore negli archivi considerati.

Sono altresì emersi aspetti negativi che l'utilizzo dei dati sanitari comporta. La qualità degli archivi è uno dei punti di maggior criticità presentati da questo strumento di ricerca epidemiologica. La difficoltà risiede nella necessità di elaborare informazioni attendibili e pertinenti, raccolte però per scopi diversi da quello di ricerca scientifica e rispondenti dunque ad esigenze non sempre coincidenti con quelle dell'epidemiologo. Mentre per finalità di sanità pubblica, infatti, si esigono archivi completi e al loro interno campi normalizzati, completi e aggiornati, per scopi epidemiologici, quali supporto alla programmazione sanitaria, monitoraggio e sorveglianza delle malattie e conduzione di studi analitici, è fondamentale conoscere la qualità degli archivi stessi, che deve poter essere controllata più che risultare ottimale. L'errore non può fisiologicamente essere eliminato, per limiti intrinseci al lavoro umano, ma deve assolutamente essere quantificato e governato, attestandosi entro soglie di accettabilità stabilite a priori.

Nel caso del SEI, nel corso dell'analisi di dati relativi alla mortalità per suicidio e omicidio nel periodo 1994-2004, insospettiti dall'assenza di casi per l'anno 1997, si è dovuti risalire all'archivio originario dei certificati di morte e riacquisire i dati relativi all'anno in questione: erroneamente non era stato fornito il campo contenente la codifica delle cause di morte esterne, quali appunto suicidio e omicidio, per il 1997. Questo episodio mostra concretamente a quali problematiche di sicurezza dei dati esponga la natura amministrativa della raccolta, spesso effettuata con scarso rigore.

Il problema maggiore, anche in questo caso, non è la presenza dell'errore in sé, quanto la sua individuazione e quantificazione.

Si rende quindi imprescindibile la produzione di indicatori di qualità degli archivi, preferibilmente istituzionalizzata per esigenze di confrontabilità e trasparenza, e la normalizzazione delle banche dati con altre fonti che possono essere ritenute gold standard.

La certificazione di qualità permetterebbe una valutazione di pertinenza delle misure ottenute al perseguimento degli scopi prefissati, una confrontabilità fra le misure interne ad uno stesso studio e fra risultati ottenuti in studi diversi e, mettendo in luce gli aspetti più lacunosi, imporrebbe un continuo miglioramento del sistema di archiviazione dei dati, con ricadute positive anche sugli altri ambiti di impiego delle informazioni.

Il problema intrinsecamente presente nello sfruttamento di basi dati amministrative per scopi epidemiologici è dato dalla divergenza di intenzionalità nella raccolta dei dati; prima di procedere al loro impiego si rende necessaria una verifica della congruenza tra le variabili disponibili e le analisi che s'intende produrre.

L'utilizzo dei dati personali sensibili, quali quelli relativi alla salute, richiede inoltre particolari cautele, che rischiano spesso di limitare le possibilità di ricerca.

Infine, se gli algoritmi per la definizione della casistica sono solitamente di agile struttura, non altrettanto si può affermare della loro costruzione: questo processo richiede, infatti, competenze multidisciplinari ed elevata accuratezza nella scelta dei codici diagnostici per le diverse fonti, affinché vi sia corrispondenza fra i record selezionati e le analisi che ci si prefigge di svolgere; è di estrema importanza, infatti, che il fenomeno di interesse sia effettivamente misurabile con i dati ottenuti.

Le problematiche concernenti questo aspetto dell'integrazione degli archivi saranno affrontate nel capitolo successivo.

In questa prima fase esplorativa sono emersi aspetti fondamentali da approfondire, ma le ampie potenzialità del SEI impongono di non scoraggiarsi di fronte a questi ostacoli, bensì di esplorarne le dinamiche per risolvere le problematiche aperte.

La prospettiva di utilizzo più interessante e innovativa di un sistema integrato di basi dati elettroniche è rappresentata dalla possibilità di definire nella popolazione gli individui colpiti da una certa malattia che sono entrati in contatto a diversi livelli con il sistema sanitario.

L'ambizioso ma non utopistico obiettivo è quello di selezionare elettronicamente, attraverso algoritmi specifici, la casistica per stimare incidenza e prevalenza delle malattie sulla popolazione e individuare coorti per condurre studi epidemiologici analitici, sempre utilizzando le innumerevoli informazioni presenti negli archivi disponibili.

Inoltre, il Sistema Epidemiologico Integrato, raggruppando in un'unica coorte la popolazione degli assistibili dell'AULSS, dalla quale è possibile estrarre specifiche coorti di malati o esposti a fattori di rischio, permette di ricostruire e seguire sistematicamente la storia naturale dello sviluppo delle patologie e i profili di malattia per sesso, età e area geografica in prospettiva eziologica, in una sorta di "longitudinalizzazione" della popolazione.

Per procedere lungo questa nuova e promettente strada si rende necessario investire energie per analizzare le problematiche emerse in modo critico, operando in sinergia con le diverse realtà che, in modo più o meno omogeneo, hanno cominciato a sperimentare l'utilizzo di archivi elettronici sanitari per finalità epidemiologiche.

Il confronto di esperienze e metodologie permette di costruire una base di riferimento comune e di condividere le strategie che si sono rivelate migliori nelle diverse situazioni: trattandosi, infatti, di attività sperimentale viene a

mancare un riferimento accreditato per la metodologia utilizzata e la valutazione dei risultati ottenuti.

Nel prossimo capitolo verrà presentato il lavoro di un gruppo nazionale che ha tentato di affrontare questi aspetti con la discussione dei risultati ottenuti.

Capitolo 3

RISULTATI: IL GRUPPO DI LAVORO NAZIONALE AIE-SISMEC

Nel 2005 si è costituito un gruppo di lavoro nazionale allo scopo di confrontare le esperienze condotte da diversi centri di ricerca epidemiologica con l'utilizzo di archivi elettronici, per sintetizzare approcci e metodologie differenti applicati allo stesso ambito, nella convinzione che l'utilizzo integrato e diversificato degli archivi elettronici contenenti dati sanitari possa divenire strumento di supporto fondamentale per stimare la frequenza delle malattie nella popolazione ad un livello territoriale dettagliato e costituisca la nuova frontiera dell'epidemiologia analitica.

Al gruppo hanno aderito l'Associazione Italiana di Epidemiologia (AIE) e la Società Italiana di Statistica Medica ed Epidemiologia Clinica (SISMEC), sostenendo fortemente come sia necessario conoscere l'avanzamento a livello nazionale di questo emergente filone di ricerca, per evitare che si intraprendano strade parallele sempre più difficilmente ricongiungibili con il passare del tempo. Solo condividendo in sinergia competenze e problematiche, la ricerca può strutturarsi in modo completo ed avanzato.

Gli aspetti da approfondire sono stati suddivisi in 3 macro-aree:

- stime di patologie di popolazione
- procedure di record linkage probabilistico fra gli archivi
- utilizzo delle basi di dati per studi di epidemiologia analitica ed individuazione delle covariate sociali

Il sottogruppo per la produzione di stime di patologie di popolazione si è posto l'obiettivo di stabilire, in via sperimentale, le regole da rispettare nell'utilizzo degli archivi elettronici, da consolidare in vista di un'eventuale istituzionalizzazione delle procedure con estendibilità a livello nazionale. Si

rendeva necessario esplorare in dettaglio i limiti e le potenzialità che si presentano per stimolare la ricerca in questo ambito.

Il sottogruppo per l'esame delle procedure di record linkage probabilistico si è proposto per la sperimentazione e la simulazione dei metodi conosciuti per l'integrazione delle basi di dati a disposizione dell'epidemiologia.

Il terzo sottogruppo intendeva studiare l'utilizzo di covariate socio-economiche individuate tramite archivi elettronici per arricchire gli studi di epidemiologia analitica.

I centri partecipanti si sono così suddivisi, sulla base del contributo di competenze che potevano mettere a disposizione:

Stime di patologie di popolazione	Università degli Studi di Padova, Dipartimento di Medicina Ambientale e Sanità Pubblica Aulss 12 Venezia (SEI) Asl Città di Milano Agenzia di Sanità Pubblica Regione Lazio Ulss 9 Treviso Ulss 4 Thiene CNR Pisa Dipartimento di Epidemiologia Asl Roma E U.O. di Epidemiologia Asl 10 Firenze ARS Toscana Epidemiologia dei Tumori C.P.O. Piemonte U.O. di Epidemiologia CSPO Asl 10 Veneto Istituto Superiore di Sanità CPA Torino Asl 5 Torino Asl Taranto
Procedure di record linkage probabilistico fra gli archivi	Università degli Studi di Milano - Bicocca, Dipartimento di Statistica

	CNR Pisa Università di Padova – Aulss 12 Venezia (SEI) Agenzia di Sanità Pubblica Regione Lazio Università di Bari – Dipartimento di Igiene ARPA Piemonte
Utilizzo delle basi di dati per studi di epidemiologia analitica e individuazione delle covariate sociali	Università degli studi di Torino – Dipartimento di Sanità Pubblica e Microbiologia ARPA Piemonte Asl Grugliasco – Torino

In questa sede verrà presentato e discusso il lavoro del sottogruppo per la produzione di stime di patologie di popolazione, mostrandone la metodologia seguita, i risultati ottenuti e le criticità emerse che richiedono ulteriore approfondimento.

La valutazione farà riferimento essenzialmente ai risultati riferiti alla popolazione degli assistiti dall’Azienda ULSS veneziana, presso la quale la sottoscritta ha effettuato un’esperienza trimestrale di stage assistendo anche alle ultime riunioni del gruppo.

Definizione degli obiettivi

Dopo aver condiviso i diversi lavori fatti dai centri partecipanti mediante l’utilizzo di archivi elettronici, sono stati stabiliti gli obiettivi di un lavoro comune, definendo quale apporto conoscitivo ma soprattutto metodologico ci si prefiggeva di fornire.

Conoscere l’occorrenza di una malattia nella popolazione, la sua distribuzione spazio-temporale, tentare di interpretare differenze territoriali e di periodo nella diffusione di determinate patologie sono questioni delle

quali è tenuta a farsi carico l'epidemiologia, innanzitutto fornendo stime di frequenza delle malattie.

Il contributo che questo gruppo si è proposto di dare riguarda la modalità di realizzazione di queste stime.

Anziché ricorrere ad indagini ad hoc per ottenere stime di incidenza o prevalenza di una determinata patologia, la proposta è quella di utilizzare i dati contenuti negli archivi informatizzati.

Si tratta essenzialmente di intercettare i casi di una malattia incrociando le informazioni di più basi dati, senza ricorrere ad interviste personali.

Questo metodo si basa su alcuni assunti:

- di norma, ogni contatto che avviene tra assistito e sistema sanitario è registrato in un apposito archivio e riporta informazioni, codificate, che possono permettere di risalire alla motivazione del contatto;
- la maggior parte delle diagnosi di malattia è effettuata sulla base dei risultati ottenuti da esami specifici e segue un processo logico semplice.

Partendo da queste considerazioni, si è delineata la possibilità di definire un caso mediante un algoritmo decisionale che ripercorre il percorso logico che porta alla diagnosi di malattia.

Questo processo è stato applicato dapprima nell'ambito dei Registri Tumori, per sostituire l'elaborazione manuale delle informazioni nei numerosi casi in cui i dati provenienti dalle diverse fonti fossero concordanti (il successo della procedura automatica si aggira attorno al 50% dei casi, raggiungendo anche il 75%). L'algoritmo rimanda la decisione all'operatore qualora, invece, le indicazioni siano discordanti e si renda inevitabile rivedere le informazioni già computerizzate o, se necessario, la cartella clinica.

Il gruppo di lavoro si è prefissato l'obiettivo di proporre algoritmi specifici che permettano di definire i casi incidenti o prevalenti di alcune patologie

molto diffuse nella popolazione, analizzando i limiti di applicabilità, i problemi di interpretazione dei risultati ed esplorando le potenzialità di questo approccio.

Le stime di frequenza prodotte sono relative al periodo 2002-2004.

Definizione degli algoritmi e delle misure di frequenza delle malattie

Per condurre il lavoro sono state scelte le seguenti 9 patologie: infarto acuto del miocardio (IMA), cardiopatia ischemica, ictus acuto, diabete, asma, broncopneumopatia cronica ostruttiva (BPCO), malattie polmonari croniche ostruttive (MPCO), cancro della mammella, cancro del colon-retto.

Gli archivi di dati utilizzati sono quelli contenenti i certificati di morte (CM), le schede di dimissione ospedaliera dei ricoveri ordinari e in regime di day hospital (SDO), la mobilità passiva (MP), i referti di anatomia patologica (AP), le esenzioni ticket (ET) e l'archivio delle prescrizioni farmaceutiche (APF).

Per ogni patologia, più centri si sono impegnati ad applicare gli algoritmi stabiliti per confrontare poi i risultati ottenuti in contesti diversi sotto il profilo della popolazione e dell'organizzazione degli archivi.

Laddove la disponibilità delle basi di dati nei diversi centri non risalisse allo stesso anno, si è stabilita univocamente una data di partenza, rappresentata dal primo anno comune di utilizzabilità, affinché nel confronto fra le stime ottenute non si dovesse tener conto del diverso apporto delle fonti nel tempo, dimensione difficilmente quantificabile e interpretabile.

Per ogni patologia si rendeva necessario definire un algoritmo che permettesse di definire un caso di malattia in modo univoco incrociando le fonti informative disponibili. Avvalendosi della consulenza di clinici specialisti delle varie patologie, sono stati strutturati gli algoritmi, aggiornati

di volta in volta a partire dalle problematiche emerse dall'analisi dei risultati precedenti.

Come misura di frequenza è stata calcolata per alcune patologie l'*incidenza annuale*, ossia la proporzione di nuovi casi di malattia nella popolazione in un anno, mentre per altre patologie la prevalenza, che misura la proporzione di malati in un preciso momento (cfr. pag. 6).

Per quanto riguarda la prevalenza, si è scelto di calcolarla in due modi differenti a seconda delle caratteristiche della malattia; si distingue così tra la *prevalenza annuale*, definita come numero di malati presenti nella popolazione nell'anno t, calcolati sulla base delle fonti riferite a quell'anno, e la *prevalenza "longitudinale"*, che considera alcune fonti (SDO, ET) per un intervallo temporale più ampio. In questo modo si costruisce una coorte di malati e la stima risulta più attendibile per quelle patologie croniche dalle quali non sia possibile guarire e per le quali gli eventi successivi al primo si manifestino solitamente con cadenza superiore ai 12 mesi.

La prevalenza viene calcolata seguendo questo secondo metodo soltanto nel caso di malattie che presentino le suddette particolari condizioni, allo scopo di fornire una stima più attendibile.

Per la stima dell'incidenza, poiché è possibile che alcune malattie si ripresentino in uno stesso individuo a distanza di tempo, si rende necessario stabilire quando due eventi debbano essere attribuiti alla stessa malattia o quando piuttosto il soggetto si ammali nuovamente di una malattia dalla quale era guarito.

Dal punto di vista della stima della frequenza della patologia questo aspetto è rilevante: nel primo caso il soggetto, pur presentando due manifestazioni della stessa malattia, contribuisce soltanto una volta al calcolo dell'incidenza; nel secondo caso, trascorso il tempo stabilito, lo stesso soggetto contribuisce due volte alla stima in periodi successivi.

Si è stabilito un tempo di 60 mesi necessario per distinguere casi prevalenti da nuovi casi incidenti: se due eventi distano meno di 5 anni vengono attribuiti alla stessa malattia, in caso contrario vengono considerati due casi incidenti distinti.

Le stime prodotte per le patologie indagate sono così suddivise:

- incidenza annuale per IMA, ictus acuto, cancro alla mammella e cancro al colon-retto;
- prevalenza annuale per asma e MPCO;
- prevalenza utilizzando longitudinalmente le fonti SDO ed ET per cardiopatia ischemica, diabete e BPCO.

Per quanto concerne la fonte SDO, come data dell'evento nel calcolo dell'incidenza viene utilizzata la data di ammissione, mentre nel calcolo della prevalenza la data di dimissione; la scelta è motivata dal fatto che per il calcolo dell'incidenza interessa il nuovo evento, quindi al momento del suo manifestarsi, mentre per il calcolo della prevalenza interessano i soggetti che permangono malati nel periodo, quindi coloro che nella diagnosi di dimissione presentano ancora la malattia.

E' stato utilizzato anche l'archivio della mobilità passiva (cfr. pag. 10) per non escludere dalla stima i casi appartenenti alla popolazione in studio che hanno usufruito di servizi presso strutture afferenti ad altre Aziende Sanitarie.

Nel calcolo dei tassi annuali, al denominatore si utilizza la popolazione al 30 giugno dell'anno, ricavata da fonte ISTAT (www.demo.istat.it); nel calcolo del tasso di prevalenza "longitudinale" la data di riferimento per la popolazione e il calcolo dell'età dei soggetti è il 1 gennaio dell'anno di stima.

Per il calcolo dei tassi standardizzati, la popolazione di riferimento è quella italiana al censimento del 2001, ottenuta da fonte ISTAT

(<http://dawinci.istat.it>); le classi di età sono quinquennali fino ad anni “85 o più”:

CLASSI DI ETÀ	Totale
Meno di 5	2 618 794
Da 5 a 9	2 679 104
Da 10 a 14	2 805 287
Da 15 a 19	2 963 629
Da 20 a 24	3 424 350
Da 25 a 29	4 246 776
Da 30 a 34	4 543 782
Da 35 a 39	4 623 588
Da 40 a 44	4 065 579
Da 45 a 49	3 739 570
Da 50 a 54	3 849 691
Da 55 a 59	3 324 773
Da 60 a 64	3 464 947
Da 65 a 69	3 079 948
Da 70 a 74	2 803 512
Da 75 a 79	2 286 776
Da 80 a 84	1 235 317
Da 85 e più	1 240 321
Totale	56 995 744

La coorte della popolazione veneziana in studio nel periodo 2002-2004 ammonta a 908 368 soggetti, dei quali 432 130 uomini e 476 239 donne, così distribuiti nei 3 anni:

	uomini	donne	totale
2002	143473	158272	301745
2003	143946	158641	302587
2004	144711	159326	304037
totale	432130	476239	908369

Per ogni patologia viene calcolato il contributo di ciascuna fonte alla definizione dei casi, distinguendo *contributo esclusivo*, determinato da soggetti identificati come malati per mezzo di quell'unica fonte, e *contributo*

assoluto, come numero di casi alla cui identificazione ha contribuito la fonte considerata.

Per costruzione, la somma dei contributi esclusivi per le diverse fonti e dei soggetti identificati dalla combinazione di più di una fonte è pari al 100% dei casi individuati; la somma dei contributi assoluti è invece $\geq 100\%$.

Per ogni patologia è contemplato uno specifico controllo di qualità delle stime prodotte; un parametro comune utilizzato come indicatore di qualità degli archivi è costituito dalla percentuale per anno, sesso e classi di età di SDO e CM che riportano come diagnosi un codice del raggruppamento “Sintomi, segni e stati morbosi mal definiti” (ICD IX codici 780-799), perché possibile indice di scarsa attendibilità dei dati.

In questa sede si fa riferimento all’ultima modifica disponibile del protocollo, aggiornato in data 28 maggio 2007. Questa versione non è necessariamente definitiva per la diffusione degli algoritmi, ma si caratterizza come punto di arrivo di questa fase di lavoro del gruppo e dunque riferimento pertinente per una valutazione critica.

Infarto acuto del miocardio (IMA)

L’infarto acuto del miocardio è definito come necrosi del tessuto muscolare cardiaco in seguito ad occlusione di una coronaria, a spasmi coronarici o a squilibri metabolici prolungati.

Per stimare l’incidenza annuale di IMA nella popolazione attraverso il solo utilizzo di archivi elettronici si è definito “caso” il primo ricovero o il decesso di un soggetto registrato nell’archivio delle SDO o dei certificati di

morte che riportasse il codice ICD IX 410¹¹, specifico della patologia, nelle seguenti combinazioni:

FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
CM	410	principale	
SDO	410	principale	esclusione ricoveri in regime di Day Hospital
SDO	410	secondaria	<p>se presente nella diagnosi principale uno dei seguenti codici:</p> <p>427.1 Tachicardia parossistica ventricolare, 427.41 Fibrillazione ventricolare, 427.42 Flutter ventricolare 427.5 Arresto cardiaco 428.1 Insufficienza del cuore sinistro (scompenso cardiaco sinistro) 429.5 Rottura di corda tendinea 429.6 Rottura di muscolo papillare 429.71 Difetto settale acquisito 429.79 Altri postumi di infarto miocardico non classificati altrove (Trombo murale acquisito, successivo a infarto miocardico) 429.81 Altre alterazioni del muscolo papillare 518.4 Edema polmonare acuto, non specificato 780.2 Sincope e collasso 785.51 Shock cardiogeno 414.10 Aneurisma cardiaco (parete) 423.0 Emopericardio</p>

Per ripulire la stima di incidenza dai soggetti prevalenti (cfr. pag. 29) si è considerato un intervallo di 60 mesi durante il quale non sia stato registrato nell'archivio SDO alcun ricovero dello stesso soggetto con codice ICD IX 410 o 412 in qualsiasi posizione di diagnosi.

¹¹ ICD IX 410: Infarto miocardico acuto

Per Venezia non è stato ancora possibile elaborare i dati di mortalità dell'ultimo trimestre del 2004; di conseguenza i tassi relativi a quest'anno sono probabilmente sottostimati.

Data l'esiguità dei casi nella popolazione giovane, la prima classe di età comprende soggetti fino a 14 anni compiuti.

Tab. 1-IMA: Riepilogo dei casi incidenti di infarto acuto del miocardio, per sesso, classi di età e anno di incidenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età						totale
		0-14	15-34	35-54	55-64	65-74	75+	
uomini	2002	0	2	53	63	93	135	346
	2003	0	2	40	77	93	127	339
	2004	0	4	50	51	80	113	298
tot		0	8	143	191	266	375	983
donne	2002	0	1	7	19	49	172	248
	2003	0	3	8	21	53	170	255
	2004	0	1	10	14	49	155	229
tot		0	5	25	54	151	497	732
totale	2002	0	3	60	82	142	307	594
	2003	0	5	48	98	146	297	594
	2004	0	5	60	65	129	268	527
tot		0	13	168	245	417	872	1.715

Tab. 2-IMA: Tassi di incidenza (valori *100 000) stimati per infarto acuto del miocardio, per sesso, classi di età e anno di incidenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età						Tasso Grezzo	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75+	pop al 30/06		
uomini	2002	0,0	5,7	121,5	299,9	582,0	1.217,5	241,2	235,2	210,2 - 260,2
	2003	0,0	5,9	90,8	365,5	573,8	1.106,7	234,8	231,2	206,1 - 256,2
	2004	0,0	12,1	112,1	243,3	485,2	955,4	205,9	197,7	174,9 - 220,5
tot								227,2	221,0	207,0 - 235,0
donne	2002	0,0	3,1	15,9	80,9	234,3	793,5	156,7	103,2	90,1 - 116,2
	2003	0,0	9,6	18,0	89,7	251,8	765,9	160,7	107,4	93,9 - 120,8
	2004	0,0	3,3	22,2	60,6	231,5	685,1	143,7	93,3	80,9 - 105,7
tot								153,7	101,2	93,8 - 108,7
totale	2002	0,0	4,5	67,8	184,3	384,9	937,0	196,8	160,1	147,2 - 173,0
	2003	0,0	7,7	53,5	220,3	391,9	882,1	196,0	158,6	145,7 - 171,4
	2004	0,0	7,8	134,8	147,5	342,6	777,9	173,3	138,8	126,8 - 150,8
tot								188,7	152,4	145,1 - 159,6

I tassi di incidenza assumono consistenza maggiore negli uomini rispetto alle donne a tutte le età e hanno un andamento fortemente crescente all'aumentare dell'età.

Nell'anno 2004 si nota una notevole diminuzione del numero di casi e un conseguente decremento del tasso di incidenza sia negli uomini che nelle donne. Osservando la distribuzione per fasce d'età, la diminuzione si riscontra soltanto dalla classe 55-64 anni; la differenza con il tasso relativo all'anno precedente non è comunque significativa, poiché gli intervalli di confidenza si intersecano.

Poiché i casi vengono identificati attraverso gli archivi elettronici, la differenza potrebbe essere il risultato di anomalie presenti nelle fonti utilizzate e non costituire un fenomeno effettivamente verificatosi nella popolazione.

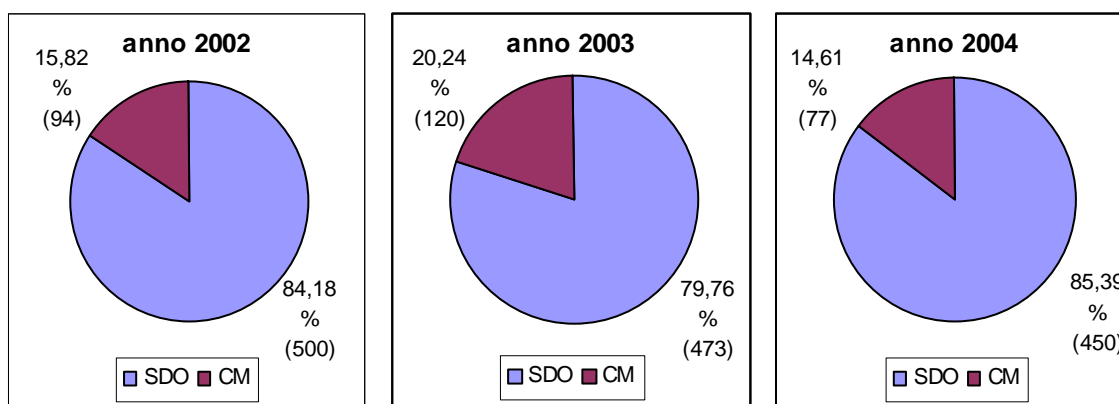
Per valutare questo aspetto si riportano i contributi delle varie fonti nella determinazione dei casi incidenti, per anni di stima e classi d'età.

Per l'infarto si tratta di contributi esclusivi, poiché l'algoritmo considera informazioni contenute nello stesso archivio nelle diverse definizioni di caso e non è contemplata una combinazione di più fonti per definire un evento.

Tab. 3-IMA: Numero assoluto e percentuale dei casi di IMA identificati da fonte SDO, CM e più fonti (contributo esclusivo), per anno e classi d'età a partire da 15-34 anni. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

fonte	classi d'età									
	15-34		35-54		55-64		65-74		75+	
	n°	%	n°	%	n°	%	n°	%	n°	%
2002										
SDO	3	100,00%	58	96,67%	69	84,15%	125	88,03%	245	79,80%
CM	0	0,00%	2	3,33%	13	15,85%	17	11,97%	62	20,20%
tot	3	100,00%	60	100,00%	82	100,00%	142	100,00%	307	100,00%
2003										
SDO	4	80,00%	40	83,33%	87	88,78%	126	86,30%	217	73,06%
CM	1	20,00%	8	16,67%	11	11,22%	20	13,79%	80	26,94%
tot	5	100,00%	48	100,00%	98	100,00%	146	100,00%	297	100,00%
2004										
SDO	4	80,00%	58	96,67%	57	87,69%	112	86,82%	219	81,72%
CM	1	20,00%	2	3,33%	8	12,31%	17	13,18%	49	18,28%
tot	5	100,00%	60	100,00%	65	100,00%	129	100,00%	268	100,00%

Grafico 1-IMA: Distribuzione percentuale del contributo esclusivo delle fonti SDO e CM, per anno di stima. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.



Dalla tabella 3 emerge come la distribuzione dei contributi delle due fonti non presenti differenze rilevanti; osservando le distribuzioni marginali, riportate anche nel grafico, l'andamento per l'anno 2004 è in linea con l'anno 2002. Non sembra dunque si possa attribuire la differenza riscontrata nei tassi del 2004 ad una diversa determinazione degli stessi: i contributi esclusivi sono molto simili nei tre anni considerati, soprattutto nelle classi d'età dove si concentra il maggior numero di casi d'infarto.

In realtà, mancando ancora i dati di mortalità dell'ultimo trimestre del 2004, il numero dei casi conoscerà verosimilmente un aumento, seppur contenuto, che porterà la stima dei tassi di incidenza più vicina a quelle degli anni precedenti.

Si può notare come circa l'80% di eventi sia individuato da fonte ospedaliera (dato medio del periodo 83%), e la tabella 4 mostra come la differenza fra i due sessi sia trascurabile.

Questo aspetto è stato considerato come indicatore di qualità degli archivi, poiché si ritiene che, se un sistema sanitario offre adeguati servizi di prevenzione e assistenza, la percentuale di soggetti deceduti per IMA non debba essere eccessiva; in caso contrario, il dato può indicare scarsa accuratezza nell'archiviazione delle informazioni.

Tab. 4-IMA: Numero assoluto e percentuale dei casi di IMA identificati da fonte SDO o CM (contributo esclusivo), per anno e sesso. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

fonte	uomini		donne		totale	
	n°	%	n°	%	n°	%
	2002					
SDO	288	83,24%	212	85,48%	500	84,18%
CM	58	16,76%	36	14,52%	94	15,82%
tot	346	100,00%	248	100,00%	594	100,00%
	2003					
SDO	278	82,25%	195	76,47%	473	79,76%
CM	60	17,75%	60	23,53%	120	20,24%
tot	338	100,00%	255	100,00%	593	100,00%
	2004					
SDO	256	85,91%	194	84,72%	450	85,39%
CM	42	14,09%	35	15,28%	77	14,61%
tot	298	100,00%	229	100,00%	527	100,00%

Nel 2003 più casi sono stati segnalati da certificato di morte; in realtà, la distribuzione dei contributi delle due fonti subisce una notevole variazione per le donne (differenza di 9 punti percentuali), mentre si mantiene pressoché costante per gli uomini; per quanto riguarda le età, la differenza è riscontrabile nelle fasce 35-54 anni e 75+.

Come si può notare dalle tabelle 1 e 2, il numero dei casi fra 2002 e 2003 si può considerare costante e il tasso standardizzato non subisce variazioni significative per entrambi i sessi.

Da queste osservazioni sembra plausibile che vi sia stato un incremento dei decessi nelle donne nel 2003 (si passa da 36 a 60 soggetti identificati da certificato di morte), non rispondente però ad un aumento dei nuovi casi di malattia, quanto piuttosto ad un differente esito della stessa; in questa sede

non è possibile indagare più a fondo questo aspetto, che meriterebbe certamente un approfondimento per capire a cosa sia imputabile questo fenomeno. A questo proposito, sarebbe interessante disporre della distribuzione del contributo delle fonti per sesso e classi d'età.

Il numero di record presenti negli archivi di mortalità a prescindere dal codice è maggiore nel 2003 rispetto all'anno precedente; da altre elaborazioni effettuate, la mortalità generale della popolazione veneziana sembra avere un incremento nel 2003 rispetto agli anni limitrofi, ma questo si riscontra soltanto nel sesso femminile per il gruppo di cause di morte per malattie del sistema circolatorio¹², comprendenti anche la patologia indagata. Confrontando l'andamento dei tassi standardizzati con gli altri centri che hanno partecipato alla produzione di stime di incidenza per IMA, Venezia riporta tassi età-specifici più bassi nei maschi; poiché l'andamento per le femmine è paragonabile a quello degli altri centri, potrebbe darsi che si possa escludere un problema di completezza degli archivi.

L'algoritmo per la stima dell'incidenza di infarto acuto del miocardio ha mostrato un funzionamento soddisfacente e i risultati sono apparsi omogenei fra i centri.

Confrontando le stime prodotte con i dati ottenuti nell'ambito del progetto *Registro Nazionale degli Eventi Coronarici e Cerebrovascolari*, pur con le dovute cautele avendo impiegato una differente metodologia e trattandosi di definizioni di malattia non sovrapponibili, sembra si possa considerare i valori sostanzialmente simili: in Friuli, il tasso di attacco per eventi coronarici non fatali è stato stimato attorno a 230 casi per 100 000 abitanti nel periodo 1998-1999¹³.

¹² Si tratta del grande gruppo di cause di morte "VII: Malattie del sistema circolatorio", comprendente i codici ICD IX da 390 a 459.

¹³ Cfr. Bibliografia: Giampaoli S. et al. (Gruppo di Coordinamento), 2004.

Ictus acuto

L'ictus è definito come episodio acuto di origine vascolare cerebrale, causato da infarto o emorragia, caratterizzato di solito da improvvisa perdita di coscienza e da emiplegia, ossia paralisi parziale o completa che interessa una metà laterale del corpo.

Il gruppo di lavoro ha rivolto l'attenzione alle forme acute di ictus, poiché le forme più lievi possono non essere registrate negli archivi utilizzati, entrando in contatto con il sistema sanitario soltanto a livello del medico di base, e questo comporterebbe la produzione di una stima distorta dell'incidenza della patologia.

Per stimare l'incidenza annuale di ictus acuto nella popolazione attraverso il solo utilizzo di archivi elettronici si è definito "caso" un soggetto, registrato nell'archivio dei certificati di morte o delle SDO, identificato come primo ricovero o decesso per mezzo dei codici 430, 431, 434 e 436¹⁴, specifici della patologia, nel modo seguente:

FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
CM	430,431,434,436	principale	
SDO	430,431,434,436	principale	esclusione ricoveri in regime di Day Hospital e SDO con anche diagnosi secondaria di 438 ¹⁵

Per ripulire la stima di incidenza dai soggetti prevalenti (cfr. pag. 29) si è considerato un intervallo di 60 mesi durante il quale non sia stato registrato nell'archivio SDO alcun ricovero dello stesso soggetto con codice ICD IX 430, 431, 434 o 436 in qualsiasi posizione di diagnosi.

¹⁴ ICD IX 430: Emorragia subaracnoidea, ICD IX 431: Emorragia cerebrale, ICD IX 434: Occlusione delle arterie cerebrali, ICD IX 436: Vasculopatie cerebrali acute, mal definite.

¹⁵ ICD IX 438: Postumi delle malattie cerebrovascolari.

Così come per l'IMA, data l'esiguità dei casi nella popolazione giovane, la prima classe di età comprende soggetti fino a 14 anni compiuti; la classe 75+ è stata ulteriormente scissa in 75-84 e 85+, per cogliere l'andamento della malattia nei più anziani.

Tab. 1-ictus: Riepilogo dei casi incidenti di ictus acuto, per sesso, classi di età e anno di incidenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età							totale
		0-14	15-34	35-54	55-64	65-74	75-84	85+	
uomini	2002	1	1	19	35	74	79	48	257
	2003	0	3	20	32	58	97	35	245
	2004	2	3	19	40	65	116	40	285
	tot	3	7	58	107	197	292	123	787
donne	2002	0	0	14	24	56	130	111	335
	2003	0	1	20	26	58	121	112	338
	2004	0	2	14	27	58	115	103	319
	tot	0	3	48	77	172	366	326	992
totale	2002	1	1	33	59	130	209	159	592
	2003	0	4	40	58	116	218	147	583
	2004	2	5	33	67	123	231	143	604
	tot	3	10	106	184	369	658	449	1779

Tab. 2-ictus: Tassi di incidenza (valori *100 000) stimati per ictus acuto, per sesso, classi di età e anno di incidenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004. Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età							Tasso Grezzo	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75-84	85+	pop al 30/06		
uomini	2002	5,9	2,9	43,6	166,6	463,1	888,3	1959,0	179,1	184,0	161,2-206,8
	2003	0,0	8,9	45,4	151,9	357,9	1037,0	1413,1	170,2	169,7	148,1-191,3
	2004	11,3	9,1	42,6	250,3	394,2	1192,7	1711,8	196,9	193,0	170,1-215,9
tot									182,1	178,6	
donne	2002	0,0	0,0	31,9	102,1	267,8	829,9	1663,6	211,7	138,2	123,3-153,2
	2003	0,0	3,2	45,0	111,2	275,6	738,7	1719,7	213,1	140,8	125,5-156,0
	2004	0,0	6,5	31,0	129,1	274,0	680,9	1726,2	200,2	132,3	117,5-147,2
tot									208,3	135,6	
totale	2002	3,1	1,5	37,7	132,6	352,4	851,1	1937,6	196,2	157,0	144,4-169,6
	2003	0,0	6,1	45,2	130,4	311,4	847,1	1851,9	192,7	153,4	140,9-165,9
	2004	5,9	7,8	36,8	181,6	326,6	867,9	1824,5	198,7	156,6	144,0-169,2
tot									195,8	153,6	

Osservando la distribuzione per età dei tassi di incidenza per ictus acuto si evidenzia una crescita notevole del tasso con l'avanzare dell'età.

L'andamento è simile nei due sessi, con diverse entità. Anche l'ictus acuto si configura come una malattia che colpisce maggiormente il sesso maschile; le differenze non sembrano essere però omogenee negli anni e per fasce d'età.

Tenendo presente l'ampiezza non omogenea delle prime 3 classi d'età, per entrambi i sessi i tassi specifici aumentano in modo consistente già a partire dalla fascia 55-64 anni, ma in modo particolare dalla successiva.

Si può notare come l'entità dei tassi per la classe d'età 75-84 anni, che nel 2002 è piuttosto simile per i due sessi, presenti una tendenza opposta nei due anni successivi: se per le donne il tasso specifico diminuisce, per gli uomini

assume valori progressivamente più elevati. Per la classe d'età successiva si registra il fenomeno opposto, con una tendenza all'aumento nelle donne che arrivano a superare gli uomini, per i quali si osserva una diminuzione consistente. La stabilità dei tassi totali specifici per queste classi d'età nasconde, dunque, tendenze forti nei due sessi che nel complesso si annullano.

Queste difformità tra i due sessi nelle età più anziane si mascherano anche per quanto riguarda tassi grezzi e standardizzati totali, che rimangono praticamente stabili nel tempo, essendo fortemente influenzati dalle età più avanzate.

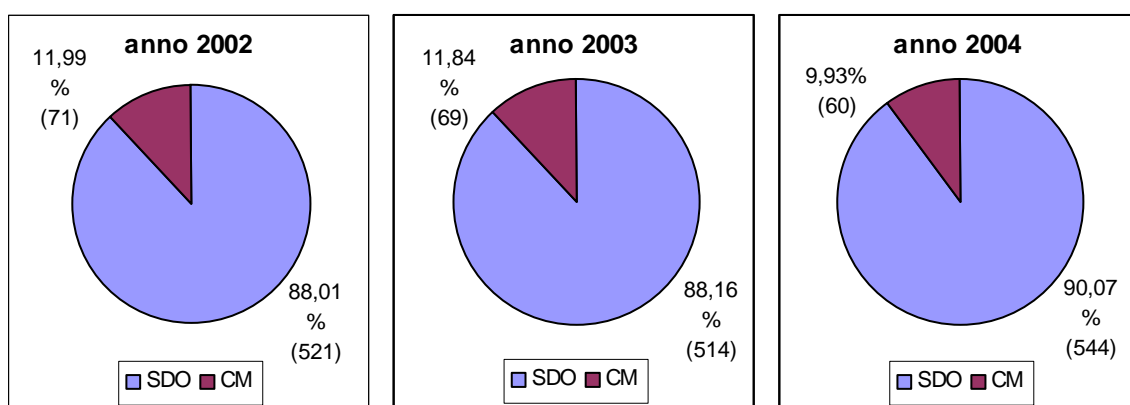
Attraverso la distribuzione dei contributi delle fonti nella determinazione dei casi si tenta di esplorare più a fondo questa dinamica; anche per l'ictus si tratta di contributo esclusivo.

Tab. 3-ictus: Numero assoluto e percentuale dei casi di ictus identificati da fonte SDO, CM e più fonti (contributo esclusivo), per anno e classi d'età a partire da 35-54 anni. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

fonte	classi d'età									
	35-54		55-64		65-74		75-84		85+	
	n°	%	n°	%	n°	%	n°	%	n°	%
	2002									
SDO	33	100,00%	58	98,31%	123	94,62%	185	88,52%	121	76,10%
CM	0	0,00%	1	1,69%	7	5,38%	24	11,48%	38	23,90%
tot	33	100,00%	59	100,00%	130	100,00%	209	100,00%	159	100,00%
	2003									
SDO	38	95,00%	55	94,83%	114	98,28%	190	87,16%	113	76,87%
CM	2	5,00%	3	5,17%	2	1,72%	28	12,84%	34	23,13%
tot	40	100,00%	58	100,00%	116	100,00%	218	100,00%	147	100,00%

	2004									
SDO	32	96,97%	63	94,03%	117	95,12%	208	90,04%	117	81,82%
CM	1	3,03%	4	5,97%	6	4,88%	23	9,96%	26	18,18%
tot	33	100,00%	67	100,00%	123	100,00%	231	100,00%	143	100,00%

Grafico 1-ictus: Distribuzione percentuale del contributo esclusivo delle fonti SDO e CM, per anno di stima. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.



Dalla distribuzione del contributo per età si nota come la percentuale di casi identificati dal certificato di morte tenda ad aumentare con l'avanzare dell'età, seppur diminuendo complessivamente nel 2004 (cfr. grafico).

Nello stesso anno, il contributo esclusivo totale delle SDO aumenta rispetto agli anni precedenti di quasi 7 punti percentuali per il sesso femminile, diminuendo invece di quasi 4 punti percentuali per i maschi.

Per approfondire questo aspetto sarebbe interessante disporre del contributo delle fonti per sesso, anno di stima ed età diviso per codice diagnostico ICD IX: la differenza osservata potrebbe essere in parte spiegata dal fatto che alcune forme di ictus colpiscono prevalentemente le donne e altre gli uomini, con differenze imputabili anche all'età del soggetto.

La tabella 4 mostra infine come nel 2004 sia aumentato il numero di casi identificati da codice 432 e 436; in particolare, la presenza eccessiva di

quest'ultimo codice, corrispondente a “vasculopatie cerebrali acute mal definite”, potrebbe essere indicatore di scarsa accuratezza diagnostica.

Tab. 4-ictus: Distribuzione assoluta e percentuale per anno dei codici ICD IX utilizzati nella definizione della casistica. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

anno	codici ICD IX									
	430		431		434		436		totale	
	n°	%	n°	%	n°	%	n°	%	n°	%
2002	26	4,39%	98	16,55%	313	52,87%	155	26,18%	592	100,00%
2003	31	5,32%	92	15,78%	321	55,06%	139	23,84%	583	100,00%
2004	18	2,98%	122	20,20%	278	46,03%	186	30,79%	604	100,00%
totale	75	4,22%	312	17,54%	912	51,26%	480	26,98%	1779	100,00%

Dal confronto con i risultati ottenuti presso gli altri centri partecipanti, Venezia riporta tassi standardizzati di ictus acuto più bassi sia negli uomini che nelle donne.

Si dovrà in futuro approfondire questo aspetto, per capire se la differenza sia dovuta ad un'effettiva minor incidenza della malattia nella popolazione veneziana o ad una distorsione introdotta dal sistema di stima, tenendo conto delle diverse abitudini terapeutiche esistenti a seconda delle aree.

Le stime fornite da altre fonti sono complessivamente più alte: ad esempio, l'incidenza grezza annuale per primo evento di ictus nei 12 mesi fra il 1 giugno 1992 e il 31 maggio 1993 per la provincia di Belluno è stata stimata in 224 casi ogni 100 000 abitanti¹⁶; si deve però tener conto della restrizione attuata dal gruppo AIE-SISMEC alla definizione di malattia, scegliendo di concentrarsi sui casi acuti di ictus. Nel confronto non si tiene in considerazione la stima del tasso standardizzato, in quanto la popolazione

¹⁶ Cfr. bibliografia: Lauria G. et al., 1995.

presa come riferimento è differente (nell'articolo citato si tratta di quella europea).

Complessivamente, dunque, i risultati ottenuti con questo algoritmo sono soddisfacenti.

Cardiopatía ischemica

La cardiopatía ischemica è una patologia che interessa l'apparato circolatorio ed è molto diffusa nella popolazione adulta ed anziana, costituendo una delle principali cause di morte nel mondo occidentale. Ne deriva l'importanza di un monitoraggio della sua diffusione nella popolazione per programmare efficaci azioni di contrasto.

Anche per questa patologia l'obiettivo del gruppo è stimarne la prevalenza negli anni 2002-2004, considerando longitudinalmente il contributo delle fonti SDO ed ET. Trattandosi di una definizione più generica, la cardiopatía comprende anche l'infarto acuto del miocardio (IMA); in questo caso, però, l'incidenza non è indagata.

L'ultima versione dell'algoritmo prevede le seguenti combinazioni di fonti:

FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
APF	Nitrati organici: C01DA*		almeno 2 prescrizioni nell'anno di stima
CM	410-414 ¹⁷	principale	nell'anno di stima
SDO	410-414	tutte	soggetti dimessi nell'anno di stima e nei 4 anni precedenti, non deceduti (per tutte le cause) al 1 gennaio dell'anno di stima

¹⁷ ICD IX 410: Infarto miocardico acuto; ICD IX 411: Altre forme acute e subacute di ischemia cardiaca; ICD IX 412: Infarto miocardico progressivo; ICD IX 413: Angina pectoris; ICD IX 414: Altre forme di ischemia cardiaca cronica.

ET	002.414		soggetti esenti nell'anno di stima e nei 3 anni precedenti, non deceduti (per tutte le cause) al 1 gennaio dell'anno di stima
----	---------	--	---

E' sufficiente una delle 4 condizioni sopra elencate per considerare un soggetto prevalente.

La versione precedente dell'algoritmo prevedeva almeno 3 prescrizioni farmaceutiche, ma per allineare la procedura di stima a quella del diabete si è preferito modificare la definizione; per ovviare ai possibili errori di codifica non ci si limita ad una sola prescrizione. I risultati ottenuti sono di seguito riportati:

Tab. 1-cardiopatìa: Riepilogo dei casi prevalenti di cardiopatìa ischemica, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età						totale
		0-14	15-34	35-54	55-64	65-74	75+	
uomini	2002	4	18	540	1302	2049	2594	6507
	2003	3	16	559	1362	2152	2745	6837
	2004	1	23	549	1265	2094	2771	6703
	tot	8	57	1648	3929	6295	8110	20047
donne	2002	1	8	138	422	1213	3633	5415
	2003	1	12	131	439	1216	3723	5522
	2004	2	19	129	405	1175	3625	5355
	tot	4	39	398	1266	3604	10981	16292
totale	2002	5	26	678	1724	3262	6227	11922
	2003	4	28	690	1801	3368	6468	12359
	2004	3	42	678	1670	3269	6396	12058
	tot	12	96	2046	5195	9899	19091	36339

Tab. 2-cardiopatia: Tassi di prevalenza (valori *100) stimati per cardiopatia ischemica per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS veneziana, 2002-2004. Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età						Tasso Grezzo pop al 30/06	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75+			
uomini	2002	0,02	0,05	1,24	6,20	12,82	23,39	4,54	4,46	4,36 - 4,56
	2003	0,02	0,05	1,27	6,46	13,28	23,92	4,75	4,60	4,50 - 4,70
	2004	0,01	0,07	1,23	6,03	12,70	23,43	4,63	4,44	4,34 - 4,54
tot								4,64	4,42	
donne	2002	0,01	0,02	0,31	1,80	5,80	16,76	3,42	2,27	2,21 - 2,33
	2003	0,01	0,04	0,29	1,88	5,78	16,77	3,48	2,29	2,23 - 2,35
	2004	0,01	0,06	0,29	1,75	5,55	16,02	3,36	2,19	2,14 - 2,25
tot								3,42	2,28	
totale	2002	0,02	0,04	0,77	3,87	8,84	19,01	3,95	3,20	3,14 - 3,25
	2003	0,01	0,04	0,78	4,05	9,04	19,21	4,08	3,27	3,22 - 3,32
	2004	0,01	0,07	0,76	3,79	8,68	18,57	3,97	3,15	3,10 - 3,20
tot								4,00	3,19	

L'andamento dei tassi età-specifici nei 3 anni di stima è sostanzialmente costante; nel 2003 la prevalenza sembra assumere valori leggermente più elevati a partire dalla classe 35-54 anni, specialmente negli uomini, mentre nel 2004 i tassi si riavvicinano ai valori del 2002. I tassi standardizzati non sono significativamente diversi al 95%.

La dinamica per età è molto marcata: la prevalenza comincia ad assumere una certa consistenza a partire dalla classe 35-54 anni, per proseguire un trend crescente che porta a valori molto elevati negli anziani, come atteso: più di un uomo su cinque oltre i 75 anni soffre di cardiopatia ischemica.

Anche questa malattia colpisce maggiormente la popolazione maschile, con un rapporto in media di quasi 2 uomini ogni donna.

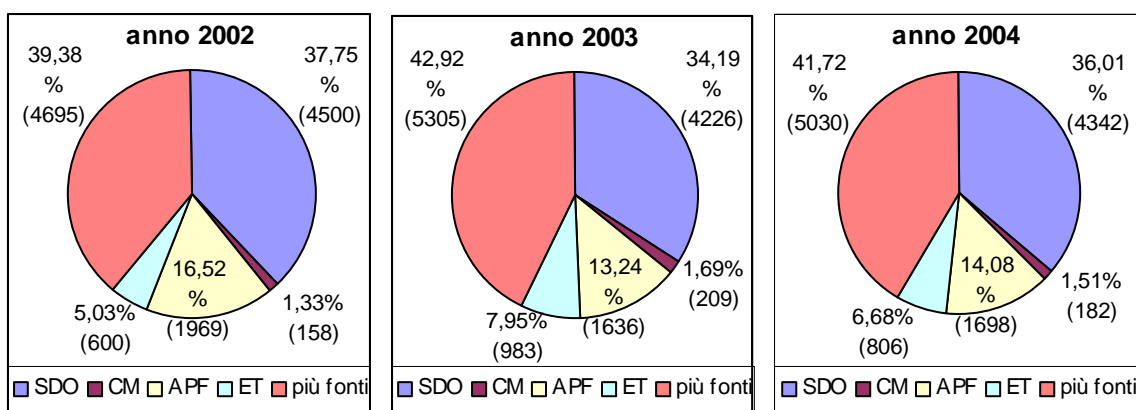
Il contributo dei vari archivi nella determinazione dei casi si distribuisce per fasce d'età come riportato in tabella 3:

Tab. 3-cardiopatia: Numero assoluto e percentuale dei casi di cardiopatia ischemica identificati da fonte SDO, CM, APF, ET e più fonti (contributo esclusivo), per anno e classi d'età. Popolazione dell'Azienda ULSS veneziana, 2002-2004.

fonte	classi d'età											
	0-14		15-34		35-54		55-64		65-74		75+	
	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%
	2002											
SDO	5	100,00%	15	57,69%	300	44,25%	658	38,17%	1277	39,15%	2245	36,05%
CM	0	0,00%	1	3,85%	4	0,59%	16	0,93%	13	0,40%	124	1,99%
APF	0	0,00%	3	11,54%	28	4,13%	175	10,15%	482	14,78%	1281	20,57%
ET	0	0,00%	0	0,00%	59	8,70%	169	9,80%	205	6,28%	167	2,68%
più fonti	0	0,00%	7	26,92%	287	42,33%	706	40,95%	1285	39,39%	2410	38,70%
tot	5	100,00%	26	100,00%	678	100,00%	1724	100,00%	3262	100,00%	6227	100,00%
	2003											
SDO	3	75,00%	11	39,29%	251	36,38%	585	32,48%	1158	34,38%	2218	34,29%
CM	0	0,00%	1	3,57%	9	1,30%	18	1,00%	27	0,80%	154	2,38%
APF	1	25,00%	2	7,14%	24	3,48%	147	8,16%	339	10,07%	1123	17,36%
ET	0	0,00%	4	14,29%	74	10,72%	236	13,10%	344	10,21%	325	5,02%
più fonti	0	0,00%	10	35,71%	332	48,12%	815	45,25%	1500	44,54%	2648	40,94%
tot	4	100,00%	28	100,00%	690	100,00%	1801	100,00%	3368	100,00%	6468	100,00%

		2004											
SDO	2	66,67%	15	35,71%	264	38,94%	610	36,53%	1197	36,62%	2254	35,24%	
CM	1	33,33%	4	9,52%	11	1,62%	17	1,02%	26	0,80%	123	1,92%	
APF	0	0,00%	6	14,29%	28	4,13%	142	8,50%	333	10,19%	1189	18,59%	
ET	0	0,00%	5	11,90%	63	9,29%	176	10,54%	291	8,90%	271	4,24%	
più fonti	0	0,00%	12	28,57%	312	46,02%	725	43,41%	1422	43,50%	2559	40,01%	
tot	3	100,00%	42	100,00%	678	100,00%	1670	100,00%	3269	100,00%	6396	100,00%	

Grafico 1-cardiopatia: Distribuzione percentuale del contributo esclusivo delle fonti SDO, CM, APF, ET e di più fonti, per anno di stima. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.



La distribuzione dei contributi esclusivi non presenta differenze importanti nei 3 anni di stima; nel complesso, circa il 40% dei casi è individuato dalla combinazione di più fonti (cfr. grafico), percentuale che assume il valore più elevato nella classe d'età 35-54 anni, per poi calare con l'avanzare dell'età, con tendenza opposta al contributo dei farmaci. I certificati di morte presentano un contributo esclusivo generalmente inferiore al 2%, fatta eccezione per la classe 15-34 anni per la quale gli episodi acuti hanno portato alla morte senza essere stati in precedenza diagnosticati con una frequenza di oltre 3 casi su 100 nel 2002 e nel 2003, superando addirittura il 9% nel 2004. Questo dato, non in linea con quelli precedenti e relativi alle

classi d'età attigue, può destare qualche preoccupazione sugli interventi sanitari ma anche qualche sospetto sulla sua attendibilità: si tenga conto che la variabilità in questa fascia è molto elevata, essendo le proporzioni calcolate su un numero esiguo di casi. Il tasso di prevalenza età-specifico nel 2004 è comunque più elevato per entrambi i sessi.

Per quanto riguarda la fonte SDO, la percentuale di casi identificati soltanto dalla scheda di dimissione ospedaliera tende a diminuire con l'età, assieme al contributo di più fonti, mentre aumentano notevolmente nell'ultima classe i soggetti catturati utilizzando i farmaci o i CM soltanto. Queste dinamiche, riscontrate in tutto il periodo indagato, riflettono verosimilmente differenti approcci di cura ai cardiopatici a seconda dell'età; potrebbe trattarsi altresì di forme differenziate di malattia, che si evolvono negli anni richiedendo trattamenti diversificati.

Confrontando le distribuzioni dei contributi delle fonti ottenute dagli altri centri, si nota come per Venezia e Treviso i farmaci contribuiscano all'identificazione di un minor numero di casi, sia come archivio singolo che nell'integrazione con le altre fonti: considerando insieme uomini e donne per l'anno 2003, il contributo esclusivo della fonte APF a Treviso rappresenta l'11,3%, a Venezia, come riportato in tabella 3, il 13,2%, mentre a Firenze e a Torino rispettivamente il 22,4% e il 23,5% dei casi è individuato solamente per mezzo di una prescrizione farmaceutica.

Questo dato è completamente in linea con quanto riportato dall'Agenzia Italiana del Farmaco a proposito dell'utilizzo dei nitrati, farmaci specifici della malattia, nelle Regioni italiane nel 2005¹⁸. Il Veneto, assieme ad altre Regioni per le quali non sono disponibili stime effettuate dal gruppo, presenta la percentuale più bassa di consumo pro-capite di nitrati e questa convergenza non fa che confermare la buona qualità delle stime prodotte mediante l'algoritmo.

¹⁸ Cfr. Bibliografia: Agenzia Italiana del Farmaco, 2006.

Asma

L'asma bronchiale è una malattia caratterizzata da un aumento dell'irritabilità dei bronchi a vari stimoli, che predispone a crisi acute di restringimento delle vie aeree (broncospasmo).

Per stimare la prevalenza annuale di asma, da calcolare ogni singolo anno, sono stati utilizzati gli archivi delle SDO, dei CM, delle prescrizioni farmaceutiche (APF) e delle esenzioni ticket (ET), secondo il seguente algoritmo:

FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
APF	<ul style="list-style-type: none">• Adrenergici per aerosol: R03A*• Adrenergici per uso sistemico: R03CC02, R03CC04, R03CK• Antagonisti dei recettori leucotrienici: R03DC01, R03DC03		almeno una prescrizione
CM	493	principale	nell'anno di stima
SDO	493	tutte	nell'anno di stima, selezionate secondo la data di dimissione
ET	007.493		nell'anno di stima (i nuovi esenti)

Viene considerato caso prevalente un soggetto che venga intercettato per avere nell'anno di stima almeno una delle 4 condizioni sopra riportate.

Per coerenza con le altre malattie, anche per l'asma vengono utilizzate le classi di età già presentate, ma come si vedrà la scelta dovrà essere ridiscussa.

Tab. 1-asma: Riepilogo dei casi prevalenti di asma, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età						totale
		0-14	15-34	35-54	55-64	65-74	75+	
uomini	2002	2328	1381	1288	827	1086	1001	7911
	2003	2251	1327	1230	765	996	991	7560
	2004	2192	1296	1309	755	1002	1006	7560
tot		6771	4004	3827	2347	3084	2998	23031
donne	2002	1596	1227	1705	1161	1308	1371	8368
	2003	1613	1098	1641	1049	1230	1328	7959
	2004	1567	1126	1640	1015	1223	1427	7998
tot		4776	3451	4986	3225	3761	4126	24325
totale	2002	3924	2608	2993	1988	2394	2372	16279
	2003	3864	2425	2871	1814	2226	2319	15519
	2004	3759	2422	2949	1770	2225	2433	15558
tot		11547	7455	8813	5572	6845	7124	47356

Tab. 2-asma: Tassi di prevalenza (valori *100) stimati per asma, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004. Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età						Tasso Grezzo pop al 30/06	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75+			
uomini	2002	13,74	3,97	2,95	3,94	6,80	9,03	5,51	5,75	5,63 - 5,87
	2003	13,04	3,92	2,79	3,63	6,15	8,64	5,25	5,45	5,33 - 5,57
	2004	12,42	3,91	2,93	3,60	6,08	8,51	5,22	5,37	5,25 - 5,49
tot								5,33	5,56	
donne	2002	10,07	3,79	3,88	4,94	6,25	6,33	5,29	5,27	5,15 - 5,38
	2003	9,98	3,50	3,69	4,48	5,84	5,98	5,02	5,01	4,90 - 5,12
	2004	9,49	3,66	3,63	4,85	5,78	6,31	5,02	4,98	4,87 - 5,09
tot								5,11	5,12	

	2002	11,97	3,88	3,42	4,47	6,49	7,24	5,39	5,48	5,40 - 5,56
totale	2003	11,56	3,72	3,24	4,08	5,97	6,89	5,13	5,21	5,13 - 5,29
	2004	11,00	3,79	3,29	4,02	5,91	7,06	5,12	5,15	5,07 - 5,23
tot								5,21	5,31	

Come si evince dalla tabella 2, l'andamento dei tassi di prevalenza per asma è irregolare rispetto all'età: si osservano i tassi età-specifici più elevati nella fascia più giovane, essi si ridimensionano poi bruscamente a partire dalla classe 15-34 anni, per tornare a salire progressivamente nelle successive. L'andamento riscontrato corrisponde a quanto noto, che l'asma cioè si presenta frequentemente nei bambini, per poi scomparire con la crescita.

La prevalenza sembra essere tendenzialmente più alta nei maschi nelle classi di età marginali e nelle femmine in quelle centrali (35-54 anni e 55-64 anni), ma la differenza fra i sessi non è molto rilevante.

Il tasso standardizzato relativo al 2002 è invece significativamente più elevato rispetto ai due anni seguenti: gli intervalli di confidenza al 95% sono disgiunti, per entrambi i sessi. In generale, la prevalenza di asma nella popolazione veneziana sembra essere in diminuzione.

Anche in questo caso si prende in considerazione la distribuzione dei contributi delle varie fonti per esplorare più approfonditamente gli andamenti osservati.

Data la struttura dell'algoritmo per la stima della prevalenza di asma, per ogni fonte si distingue il contributo assoluto, ovvero il numero di casi identificati *anche* dalla fonte considerata, dal contributo esclusivo, ovvero il numero di casi individuati *soltanto* grazie alla fonte considerata (cfr. pag. 30). Nelle tabelle seguenti si riporta solo il contributo esclusivo, essendo quello assoluto non disponibile per sesso e fasce d'età.

Tab. 3-asma: Numero assoluto e percentuale dei casi di asma identificati da fonte SDO, CM, APF, ET e più fonti (contributo esclusivo), per anno e classi d'età. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

fonte	classi d'età													
	0-14		15-34		35-54		55-64		65-74		75+		tot	
	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%
	2002													
SDO	12	0,31%	4	0,15%	6	0,20%	2	0,10%	3	0,13%	10	0,42%	37	0,23%
CM	0	0,00%	0	0,00%	0	0,00%	0	0,00%	0	0,00%	1	0,04%	1	0,01%
APF	3771	96,10%	2463	94,44%	2882	96,29%	1936	97,38%	2377	99,29%	2343	98,78%	15772	96,89%
ET	14	0,36%	18	0,69%	8	0,27%	4	0,20%	0	0,00%	1	0,04%	45	0,28%
più fonti	127	3,24%	123	4,72%	97	3,24%	46	2,31%	14	0,58%	17	0,72%	424	2,60%
tot	3924	100,00%	2608	100,00%	2993	100,00%	1988	100,00%	2394	100,00%	2372	100,00%	16279	100,00%
	2003													
SDO	17	0,44%	9	0,37%	3	0,10%	2	0,11%	5	0,22%	2	0,09%	38	0,24%
CM	0	0,00%	0	0,00%	0	0,00%	0	0,00%	0	0,00%	1	0,04%	1	0,01%
APF	3680	95,24%	2258	93,11%	2707	94,29%	1721	94,87%	2129	95,64%	2236	96,42%	14731	94,92%
ET	20	0,52%	24	0,99%	16	0,56%	5	0,28%	5	0,22%	7	0,30%	77	0,50%
più fonti	147	3,80%	134	5,53%	145	5,05%	86	4,74%	87	3,91%	73	3,15%	672	4,33%
tot	3864	100,00%	2425	100,00%	2871	100,00%	1814	100,00%	2226	100,00%	2319	100,00%	15519	100,00%
	2004													
SDO	12	0,32%	7	0,29%	6	0,20%	4	0,23%	5	0,22%	9	0,37%	43	0,28%
CM	0	0,00%	0	0,00%	1	0,03%	0	0,00%	0	0,00%	1	0,04%	2	0,01%
APF	3577	95,16%	2262	93,39%	2813	95,39%	1703	96,21%	2170	97,53%	2378	97,74%	14903	95,79%
ET	11	0,29%	22	0,91%	12	0,41%	4	0,23%	2	0,09%	3	0,12%	54	0,35%
più fonti	159	4,23%	131	5,41%	117	3,97%	59	3,33%	48	2,16%	42	1,73%	556	3,57%
tot	3759	100,00%	2422	100,00%	2949	100,00%	1770	100,00%	2225	100,00%	2433	100,00%	15558	100,00%

La percentuale di casi individuati per mezzo delle prescrizioni farmaceutiche supera sempre il 90%, raggiungendo nel 2002 quasi il 100% dei soggetti anziani.

Tra le altre, solamente la combinazione di più fonti raggiunge percentuali non trascurabili, tendenzialmente decrescenti con l'avanzare dell'età a partire dai 15-34 anni. Le altre fonti non arrivano mai a contribuire nella definizione dell'1% dei casi.

Il dato sui contributi assoluti, non riportato, mostra percentuali ancora più elevate per i farmaci (oltre 99%) e superiori anche per SDO ed ET, ad indicare che nella quasi totalità dei casi sono registrate prescrizioni farmaceutiche a carico dei soggetti definiti asmatici, associate talvolta ad esenzione ticket e più raramente ad ospedalizzazione.

I dati riportati indicano come nella terapia del soggetto asmatico abbia sicuramente un ruolo determinante l'assunzione dei farmaci, e questo è certamente supportato dalle indicazioni terapeutiche seguite dai medici. Le percentuali elevate di soggetti definiti asmatici con una sola prescrizione, senza interazione con altre fonti (oltre 70% nei bambini, per arrivare al 43% negli ultrasessantacinquenni), fa sorgere un problema fondamentale: i farmaci inseriti nell'algoritmo sono sempre indicatori di asma? Con questo algoritmo si includono soggetti in terapia di mantenimento e non realmente malati?

In particolare, le elevate percentuali di asmatici nei bambini veneziani devono far riflettere, tanto più che non trovano corrispondenza nei dati elaborati dagli altri centri, per i quali i tassi in questa fascia di età non superano l'8%.

Di frequente i bambini presentano disturbi respiratori di scarso rilievo, e i pediatri tendono a prescrivere con facilità farmaci lievi, utilizzati anche nella cura di soggetti asmatici; si è tentato di condurre un'analisi più approfondita

del fenomeno investigando su questo aspetto, cercando di discriminare soggetti realmente asmatici e soggetti in cura “cautelativa”.

A tal proposito è stata rappresentato l’andamento del numero di prescrizioni nella fascia 0-14 anni e per il totale delle età, riscontrando una tendenza molto simile ma di diversa intensità. Nei bambini, l’80% delle prescrizioni è unica, mentre sul totale della popolazione ci si ferma al 60%; seguono le 2 prescrizioni con circa il 17%, poi la percentuale diminuisce progressivamente in entrambi i casi.

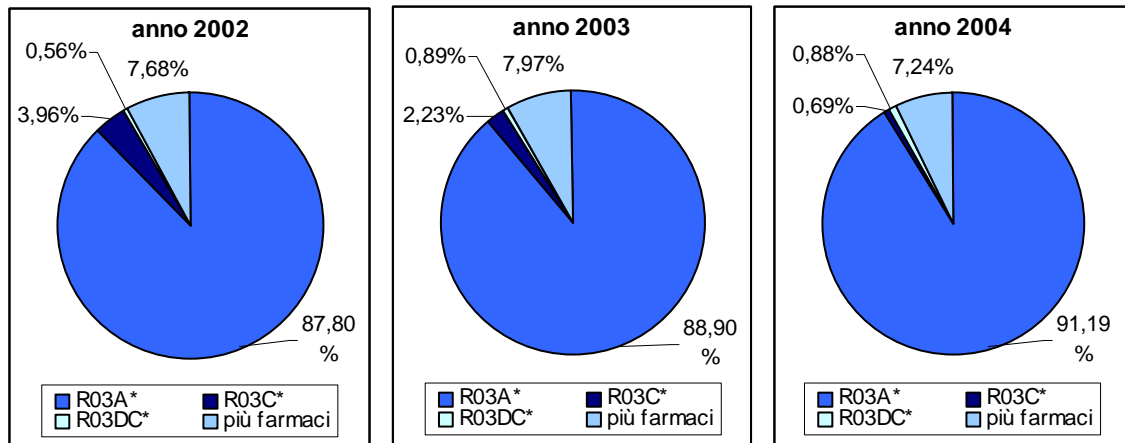
Si è anche valutata l’interazione con un altro tipo di farmaci: ai soggetti che soffrono di disturbi respiratori vengono spesso somministrati steroidi topici (cortisonici), associati ai farmaci specifici. La percentuale di questi soggetti tra i casi prevalenti individuati da una sola prescrizione nel 2003 si attesta per Venezia al 47,54%, dato in linea con quanto atteso, che rafforza almeno in parte la diagnosi di asma.

Infine, il contributo delle singole categorie farmaceutiche, alcune più discriminanti di altre, si presenta come segue:

Tab. 4-asma: Contributo esclusivo per tipologia di farmaco, per anno. La percentuale è calcolata sul contributo assoluto della fonte APF. Popolazione dell’Azienda ULSS 12 veneziana, 2002-2004.

anno	codici ATC									
	R03A*		R03C*		R03DC*		più farmaci		totale	
	n°	%	n°	%	n°	%	n°	%	n°	%
2002	14219	87,80%	642	3,96%	91	0,56%	1243	7,68%	16195	100,00%
2003	13693	88,90%	344	2,23%	137	0,89%	1228	7,97%	15402	100,00%
2004	14095	91,19%	107	0,69%	136	0,88%	1119	7,24%	15457	100,00%
totale	42007	89,27%	1093	2,32%	364	0,77%	3590	7,63%	47054	100,00%

Grafico 1-asma: Distribuzione percentuale del contributo esclusivo per tipologia di farmaco, per anno di stima, calcolata sul contributo assoluto della fonte APF. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.



Gli adrenergici per aerosol (R03A*) sono i farmaci meno discriminanti e più suscettibili di introdurre una sovrastima dei casi, più probabilmente nella realtà infantile a causa delle dinamiche riferite a pagina 55 e come si vede contribuiscono in modo nettamente maggiore rispetto alle altre tipologie.

La percentuale di soggetti che prendono più tipi di farmaci, e dunque più verosimilmente malati di asma, è comunque rilevante.

L'algoritmo per la stima della prevalenza dell'asma ha rivelato complessivamente un buon funzionamento; si tratta di approfondire l'analisi, specialmente nella fascia d'età 0-14 anni, anche suddividendola ulteriormente, per meglio comprendere le dinamiche della malattia e affinare i metodi per individuarla.

Broncopneumopatia cronica ostruttiva (BPCO)

Nell'ambito delle malattie dell'apparato respiratorio si è tentato di stimare la prevalenza annuale di BPCO, forma di bronchite cronica particolarmente diffusa fra i fumatori di sigarette e gli abitanti di zone fortemente inquinate,

che comporta tosse persistente con produzione di escreato per periodi prolungati.

Essa costituisce la più comune patologia polmonare negli uomini ed è considerata un fattore predisponente al carcinoma polmonare. L'importanza di un suo monitoraggio nella popolazione è dunque evidente, e per questo motivo il gruppo si è posto un secondo obiettivo di costituzione di una coorte di soggetti affetti da BPCO.

Per questa malattia, la stima della prevalenza di avvale del contributo longitudinale della fonte SDO, ovvero considera soggetti prevalenti i dimessi nell'anno x di stima e nei 4 anni precedenti, non deceduti (per qualsiasi causa di morte) al 1 gennaio dell'anno x. Essendo cronica, la BPCO permane infatti nel soggetto anche a distanza di anni da una fase più acuta, che ha richiesto il ricorso all'ospedalizzazione.

L'algoritmo utilizzato è il seguente:

FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
CM	490-492, 494, 496 ¹⁹	principale	decessi nell'anno di stima
SDO	490-492, 494, 496	tutte	soggetti dimessi nell'anno di stima e nei 4 anni precedenti, non deceduti (per tutte le cause) al 1 gennaio dell'anno di stima

Per la costruzione della coorte, vengono considerate le SDO con diagnosi principale di dimissione per codici 490-494, 494, 496, specifici della malattia, o le diagnosi secondarie se presente in posizione principale un codice di insufficienza respiratoria (ICD IX 518.5, 518.8, 786.0²⁰).

¹⁹ ICD IX 490: Bronchite non specificata se acuta o cronica; ICD IX 491: Bronchite cronica; ICD IX 492: Enfisema; ICD IX 494: Bronchiectasia; ICD IX 496: Ostruzioni croniche delle vie respiratorie non classificate altrove.

²⁰ ICD IX 518.5: Insufficienza polmonare successiva a trauma o a intervento chirurgico; ICD IX 518.8: Altre malattie del polmone, non classificate altrove; ICD IX 786.0: Dispnea e anomalie respiratorie.

Caso prevalente è un soggetto che riporti in almeno una SDO o nel CM i codici sopra elencati.

La BPCO è una malattia che non colpisce i giovani, quindi eventuali casi attribuiti a soggetti di età inferiore a 15 anni sono verosimilmente diagnosi errate.

Tab. 1-BPCO: Riepilogo dei casi prevalenti di BPCO, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età							totale
		0-14	15-34	35-54	55-64	65-74	75-84	85+	
uomini	2002	3	9	28	69	225	265	127	726
	2003	5	12	27	61	203	337	135	780
	2004	2	8	21	64	183	287	98	663
	2004*	20	34	84	197	514	648	180	1677
donne	2002	5	3	18	43	114	222	189	594
	2003	3	2	12	45	137	259	164	622
	2004	1	4	16	30	123	245	187	606
	2004*	10	12	49	116	349	562	356	1454
totale	2002	8	12	46	112	339	487	316	1320
	2003	8	14	39	106	340	596	299	1402
	2004	3	12	37	94	306	532	285	1269
	2004*	30	46	133	313	863	1210	536	3131

* casi prevalenti individuati con utilizzo "longitudinale" (5 anni) della fonte SDO (cfr. pag. 58)

Tab. 2-BPCO: Tassi di prevalenza (valori *100) stimati per BPCO, per sesso, classi di età e anno di prevalenza. Popolazione dell’Azienda ULSS 12 veneziana, 2002-2004. Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età							Tasso Grezzo	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75-84	85+	pop al 30/06		
uomini	2002	0,02	0,03	0,06	0,33	1,41	2,98	5,79	0,51	0,52	0,48 - 0,56
	2003	0,03	0,04	0,06	0,29	1,25	3,60	6,36	0,54	0,55	0,52 - 0,59
	2004	0,01	0,02	0,05	0,31	1,11	2,95	4,66	0,46	0,46	0,42 - 0,49
	2004*	0,11	0,10	0,19	0,94	3,15	6,76	8,67	1,16	1,14	1,09 - 1,20
donne	2002	0,03	0,01	0,04	0,18	0,55	1,42	3,14	0,38	0,25	0,23 - 0,27
	2003	0,02	0,01	0,03	0,19	0,65	1,58	2,82	0,39	0,26	0,24 - 0,28
	2004	0,01	0,01	0,04	0,13	0,58	1,45	3,26	0,38	0,24	0,22 - 0,26
	2004*	0,06	0,04	0,11	0,50	1,65	3,36	6,24	0,91	0,62	0,59 - 0,65
totale	2002	0,02	0,02	0,05	0,25	0,92	1,98	3,85	0,44	0,35	0,33 - 0,37
	2003	0,02	0,02	0,04	0,24	0,91	2,32	3,77	0,46	0,37	0,35 - 0,39
	2004	0,01	0,02	0,04	0,21	0,81	2,00	3,64	0,42	0,33	0,31 - 0,34
	2004*	0,09	0,07	0,15	0,71	2,30	4,60	6,89	1,03	0,83	0,80 - 0,86

* tassi di prevalenza stimati con utilizzo “longitudinale” (5 anni) della fonte SDO (cfr. pag. 58)

I risultati ottenuti sono conformi a quanto noto in medicina sulla diffusione della malattia, più presente nel sesso maschile, come riportato nella definizione, e al crescere dell’età. Tutti i tassi standardizzati stimati sono doppi nella popolazione maschile.

La tabella 2 mostra come la prevalenza di BPCO sia trascurabile nelle prime due classi di età e i tassi comincino a crescere dalla classe 35-54 anni, specialmente per quanto riguarda gli uomini.

Nei 3 anni di stima l'andamento della malattia sembra piuttosto costante nelle fasce d'età; è significativa la diminuzione dei tassi standardizzati di 2003 e 2004 che si riscontra negli uomini, con ricaduta anche sul totale, essendo diminuiti i tassi per le classi con peso maggiore.

Anche in questo caso, però, i dati di mortalità relativi al 2004 non sono completi, e questa tendenza potrebbe nascondere una sottostima dei casi. Si esplorerà meglio il fenomeno osservando la distribuzione dei contributi per fonte.

Per il 2004 sono stati calcolati i tassi età-specifici con il contributo quinquennale delle SDO, che accresce ovviamente le stime (più che raddoppiano), considerando anche i viventi al 1 gennaio che nei 4 anni precedenti hanno avuto un ricovero per BPCO.

Poiché i casi di malattia sono pochissimi (meno del 2%) e di scarsa attendibilità nella popolazione giovane, è opportuno non standardizzare per tutte le età, evitando così di attribuire peso ai casi incerti e di diluire il tasso. Si è deciso di troncare a 35 anni.

Tab. 3-BPCO: Tassi di prevalenza (valori *100) stimati per BPCO, per sesso e anno di stima. Popolazione dell'Azienda ULSS 12 veneziana in età ≥ 35 anni, 2002-2004.

Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	Numero		Tasso	Tasso	Tasso Stand.	I.C. 95%	Tasso	I.C. 95%*
		casi	casi*	Grezzo pop al 30/06	Grezzo*			Stand.*	
uomini	2002	714	1734	0,78	1,90	0,87	0,80 - 0,93	2,06	1,96 - 2,15
	2003	763	1696	0,82	1,84	0,91	0,85 - 0,98	1,98	1,89 - 2,07
	2004	653	1623	0,70	1,74	0,75	0,70 - 0,81	1,86	1,77 - 1,95

donne	2002	586	1459	0,53	1,33	0,41	0,38 - 0,45	1,07	1,02 - 1,13
	2003	617	1452	0,56	1,32	0,43	0,40 - 0,46	1,05	0,99 - 1,10
	2004	601	1432	0,54	1,28	0,41	0,37 - 0,44	1,01	0,96 - 1,07
totale	2002	1300	3193	0,64	1,58	0,58	0,55 - 0,61	1,46	1,41 - 1,51
	2003	1380	3148	0,68	1,55	0,60	0,57 - 0,64	1,41	1,36 - 1,46
	2004	1254	3055	0,61	1,49	0,54	0,51 - 0,57	1,35	1,30 - 1,39

* tassi di prevalenza stimati con utilizzo "longitudinale" (5 anni) della fonte SDO (cfr. pag. 58)

In effetti, restringendo la fascia d'età si apprezza maggiormente l'entità della diffusione della malattia.

Nei 3 anni la tendenza dei tassi calcolati con apporto longitudinale della fonte SDO è di diminuzione: l'andamento è significativo per gli uomini e per il totale, essendo preponderante il peso del sesso maschile.

Questi tassi assumono valori alti più del doppio rispetto a quelli semplici.

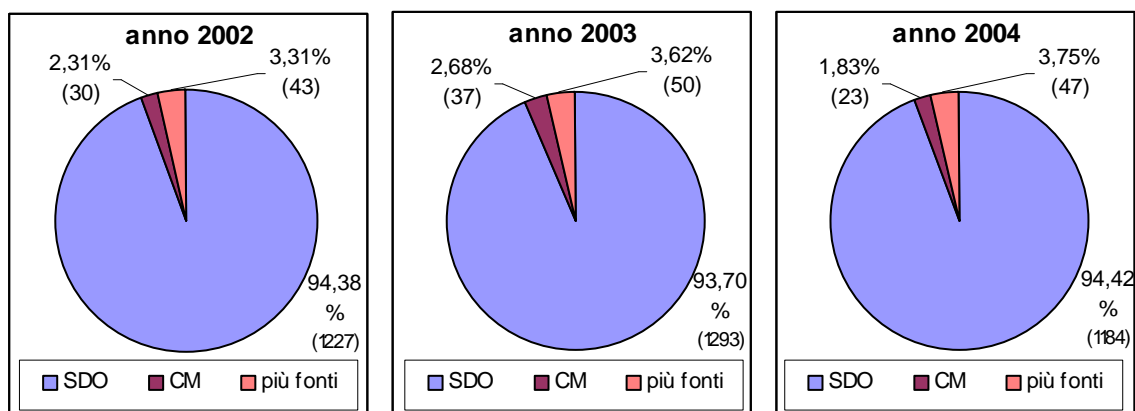
Si passa ora ad esplorare il contributo delle fonti per anno e classi d'età; per entrambe le fonti si fa riferimento agli eventi avvenuti nell'anno di stima.

Anche in questo caso si rivolge l'attenzione agli ultra trentacinquenni; dei 57 casi riscontrati in soggetti più giovani, solo uno è stato identificato esclusivamente da CM, mentre per tutti i rimanenti ha contribuito soltanto la fonte SDO.

Tab. 4-BPCO: Numero assoluto e percentuale dei casi di BPCO identificati da fonte SDO, CM e più fonti (contributo esclusivo), per anno e classi d'età. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

fonte	classi d'età									
	35-54		55-64		65-74		75-84		85+	
	n°	%	n°	%	n°	%	n°	%	n°	%
2002										
SDO	46	100,00%	110	98,21%	329	97,05%	459	94,25%	283	89,56%
CM	0	0,00%	0	0,00%	6	1,77%	10	2,05%	14	4,43%
più fonti	0	0,00%	2	1,79%	4	1,18%	18	3,70%	19	6,01%
tot	46	100,00%	112	100,00%	339	100,00%	487	100,00%	316	100,00%
2003										
SDO	38	97,44%	103	97,17%	327	96,18%	556	93,29%	269	89,97%
CM	1	2,56%	2	1,89%	2	0,59%	16	2,68%	16	5,35%
più fonti	0	0,00%	1	0,94%	11	3,24%	24	4,03%	14	4,68%
tot	39	100,00%	106	100,00%	340	100,00%	596	100,00%	299	100,00%
2004										
SDO	36	97,30%	92	97,87%	297	97,06%	492	92,48%	267	93,68%
CM	1	2,70%	0	0,00%	4	1,31%	11	2,07%	7	2,46%
più fonti	0	0,00%	2	2,13%	5	1,63%	29	5,45%	11	3,86%
tot	37	100,00%	94	100,00%	306	100,00%	532	100,00%	285	100,00%

Grafico 1-BPCO: Distribuzione percentuale del contributo esclusivo delle fonti SDO, CM e di più fonti, per anno di stima. Popolazione dell'Azienda ULSS 12 veneziana in età 35 anni o più, 2002-2004.



La fonte che contribuisce in modo preponderante è quella ospedaliera, con tendenza a decrescere all'aumentare dell'età; andamento opposto ha invece il contributo, seppur esiguo, dei CM. Nelle due classi più anziane la combinazione delle due fonti non scende mai sotto il 3,7%, ad indicare che una quota non trascurabile di malati che riceve cure per BPCO presso strutture ospedaliere in meno di 12 mesi giunge al decesso per la stessa causa.

Nei 3 anni la percentuale di soggetti ultraottantacinquenni identificati esclusivamente dalla SDO cresce, diminuendo il contributo multiplo: queste tendenze potrebbero indicare un progressivo miglioramento dell'esito delle cure offerte ai malati anziani di BPCO.

Nella classe attigua, però, si osserva la tendenza opposta: come spiegare queste differenze emerse in classi di età così vicine, peraltro non riscontrate nelle stime di prevalenza? Sarebbe interessante disporre della distribuzione del contributo per sesso e classi d'età; complessivamente, l'andamento del contributo esclusivo delle SDO per i maschi è in lieve diminuzione (1 punto percentuale nei 3 anni), mentre per le donne è irregolare, rimanendo su valori leggermente più bassi, poiché nella popolazione femminile è più elevato il contributo di più fonti.

Il contributo dei singoli codici ICD è diversificato a seconda delle età, essendo più presenti nelle cause di morte per gli anziani e nelle SDO per i giovani codici poco specifici (rispettivamente 496 e 494).

Da quanto emerso, il contributo delle varie fonti non pare risenta degli andamenti dei tassi. Questo aspetto è rassicurante, poiché indica che l'algoritmo non sembra introdurre distorsioni che vadano ad influenzare le stime.

In generale, i dati ottenuti nei centri sono abbastanza uniformi, tenendo conto delle diversità che effettivamente possono esserci nella diffusione delle malattie sul territorio.

Confrontando i tassi con quelli ottenuti per bronchite/enfisema dall'Indagine sullo stato di salute condotta dall'ISTAT per gli anni 1999-2000²¹, le stime del gruppo sono di circa quattro volte inferiori. Bisogna tenere in considerazione che la stima ISTAT è ottenuta mediante questionario rivolto al soggetto ed è quindi esposta a distorsione, dovuta tra l'altro alla soggettività della percezione del proprio stato di salute e alla scarsa affidabilità della memoria individuale nel datare un evento e nel distinguere fra diverse diagnosi di malattia piuttosto specifiche.

La stessa indagine fornisce la percentuale di fumatori in età superiore a 15 anni; essendo il fumo uno dei principali fattori di rischio, è verosimile attendersi una prevalenza di BPCO maggiore in zone con elevata proporzione di fumatori. Ebbene, in questo caso i tassi longitudinali stimati dal gruppo sembrano essere più in linea di quelli ISTAT, associando il valore più elevato a Roma e quello più basso a Venezia, dove si riscontrano rispettivamente la percentuale maggiore e minore di fumatori rispetto agli altri centri collaboratori.

²¹ Cfr. Bibliografia: ISTAT, 2002

Considerando nell'algoritmo anche l'apporto longitudinale di 5 anni della fonte ET con codice 024.518.81²², la stima si avvicina maggiormente a quella ISTAT, catturando per Venezia circa un 5% di casi aggiuntivi non identificati altrimenti.

Il metodo di selezione dei soggetti introduce verosimilmente fattori di sottostima del fenomeno: data la tipologia della malattia, spesso non viene indicata come causa principale di morte, ma specificata in quelle secondarie, mentre nell'algoritmo si include soltanto la prima diagnosi.

Inoltre, incrociando i soggetti individuati come malati di BPCO con quelli definiti asmatici, sono stati trovati numerosi casi, in modo particolare fra i giovani, risultati prevalenti per entrambe le malattie: questo dato potrebbe indicare la scarsa specificità degli algoritmi utilizzati.

La distinzione tra le due patologie mediante archivi elettronici non è agevole: sovrapposizione dei profili di malattia e utilizzo degli stessi farmaci, seppur con scopi differenti, complicano la situazione.

Per meglio cogliere le difficoltà che si presentano si propone un breve excursus sulle due patologie.

Nella sezione dedicata all'asma si è osservata una percentuale elevatissima di casi identificati dai farmaci, che vengono diffusamente impiegati per controllare le crisi, mentre questo archivio non è stato utilizzato per la selezione dei malati di BPCO.

In realtà, anche nel caso della BPCO la terapia farmacologica può essere utilizzata, ma non si rivela efficace nel contrastare l'esito della malattia, permettendo soltanto di prevenirne e controllarne i sintomi. A parte la categoria R03DC01-03 (antagonisti dei recettori leucotrienici), tutti i farmaci somministrati ai malati di asma possono essere utilizzati nel malato di BPCO.

²² Codice esenzione 024.518.81: Insufficienza respiratoria (cronica).

La distinzione di fondo fra le due malattie riguarda l'esito a lungo termine: mentre l'asma è completamente reversibile, sia spontaneamente che in seguito ad appropriato trattamento, la BPCO è praticamente irreversibile e può peggiorare al persistere dell'esposizione alle cause (in primo luogo fumo ed inquinamento).

Il fatto che l'asma trascurato possa portare all'insorgenza di BPCO rende chiara la stretta relazione esistente fra le due patologie.

Gli aspetti che le differenziano possono essere così schematizzati:

VARIABILE	ASMA	BPCO
età di insorgenza	infantile, giovane-adulta	adulta
sintomatologia	variabile a seconda dei giorni e delle ore	lentamente progressiva
elementi concomitanti	allergia, rinite e/o eczema, storia familiare di asma	lunga storia di fumo
riduzione del flusso aereo espiratorio	ampiamente reversibile	ampiamente irreversibile

Il problema che si presenta è come distinguere le due malattie selezionando i soggetti attraverso l'incrocio di archivi elettronici, utilizzando quindi soltanto le informazioni ivi contenute. Quali caratteristiche specifiche sono rintracciabili nelle fonti disponibili?

L'unico aspetto che, accanto ai codici diagnostici specifici, si può controllare negli archivi è l'età dei soggetti classificati come malati dell'una o dell'altra malattia.

Come si è osservato dai risultati ottenuti, si riscontra maggiore prevalenza di asma nei bambini di età inferiore ai 15 anni e di BPCO negli adulti a partire

dai 35 anni. Questo risultato è certamente in linea con quanto atteso; al contempo, però, vi è una porzione di malati identificati sia come asmatici che come affetti da BPCO. Si tratta di capire se le due malattie siano effettivamente concomitanti in questi soggetti, ovvero se vi sia una sovrapposizione forzata dagli algoritmi, che non hanno un sufficiente potere discriminante fra le due patologie.

Sarà importante nel futuro sperimentare altre combinazioni, per riuscire a catturare i casi di BPCO che sembrano sfuggire con il metodo finora proposto.

Malattie polmonari croniche ostruttive (MPCO)

Per controllare i risultati ottenuti dalle stime di prevalenza di asma e BPCO, si è scelto di stimare la prevalenza annuale di malattie polmonari croniche ostruttive (MPCO), definizione più ampia che include anche le due malattie respiratorie già presentate.

La struttura dell'algoritmo, simile ai due presentati a pag. 51 e 58, è la seguente:

FONTI	CODICI	DIAGNOSI	PARTICOLARITA'
APF	<ul style="list-style-type: none"> • Adrenergici per aerosol: R03A* • Adrenergici per uso sistemico: R03CC02, R03CC04, R03CK • Anticolinergici: R03BB01, R03BB02 • Derivati xantinici: R03DA01, R03DA04, R03DA05, R03DA08, R03DA11, R03DA49 		almeno una prescrizione

	<ul style="list-style-type: none"> • Antagonisti dei recettori leucotrienici: R03DC01, R03DC03 		
CM	490-493, 494, 496	principale	
SDO	490-493, 494, 496	tutte	selezionate secondo la data di dimissione
ET	007 (asma) e/o 024 (insufficienza respiratoria cronica)		nell'anno di stima (i nuovi esenti)

Se ad un soggetto viene riconosciuta almeno una delle 4 condizioni elencate nell'anno di stima, si tratta di un caso prevalente.

Tab. 1-MPCO: Riepilogo dei casi prevalenti di MPCO, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età						totale
		0-14	15-34	35-54	55-64	65-74	75+	
uomini	2002	2342	1399	1333	901	1270	1331	8576
	2003	2264	1361	1289	840	1184	1399	8337
	2004	2205	1318	1360	841	1192	1369	8285
	tot	6811	4078	3982	2582	3646	4099	25198
donne	2002	1604	1236	1753	1226	1430	1711	8960
	2003	1635	1112	1697	1129	1379	1752	8704
	2004	1570	1148	1708	1076	1395	1885	8782
	tot	4809	3496	5158	3431	4204	5348	26446
totale	2002	3946	2635	3086	2127	2700	3042	17536
	2003	3899	2473	2986	1969	2563	3151	17041
	2004	3775	2466	3068	1917	2587	3254	17067
	tot	11620	7574	9140	6013	7850	9447	51644

Tab. 2-MPCO: Tassi di prevalenza (valori *100) stimati per MPCO, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004. Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età						Tasso Grezzo	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75+	pop al 30/06		
uomini	2002	13,82	4,02	3,06	4,29	7,95	12,00	5,98	6,23	6,10 - 6,35
	2003	13,12	4,02	2,92	3,99	7,31	12,19	5,79	6,00	5,87 - 6,12
	2004	12,49	3,98	3,05	4,01	7,23	11,57	5,73	5,86	5,74 - 5,99
tot								5,83	6,05	
donne	2002	10,13	3,81	3,99	5,22	6,84	7,89	5,66	5,53	5,42 - 5,65
	2003	10,12	3,54	3,82	4,82	6,55	7,89	5,49	5,35	5,23 - 5,46
	2004	9,51	3,73	3,78	4,65	6,59	8,33	5,51	5,32	5,21 - 5,43
tot								5,55	5,44	
totale	2002	12,03	3,92	3,49	4,78	7,32	9,28	5,81	5,83	5,74 - 5,91
	2003	11,67	3,79	3,33	4,43	6,88	9,36	5,63	5,62	5,54 - 5,70
	2004	11,05	3,86	6,89	4,35	6,87	9,45	5,61	5,56	5,47 - 5,64
tot								5,69	5,70	

Essendo le due malattie polmonari più diffuse, ci si aspetta che asma e BPCO influenzino la distribuzione dei tassi di prevalenza per MPCO; i valori assoluti contenuti nella tabella 1 sono in linea con quanto atteso.

Osservando i tassi riportati nella tabella 2 si nota come siano più elevati nei soggetti giovani e tra i 15 e i 54 anni nelle donne rispetto agli uomini, come nel caso dell'asma, mentre all'avanzare dell'età la distribuzione abbia un andamento simile a quello riscontrato per la BPCO.

La prevalenza di malattie polmonari negli anziani, specialmente tra i maschi, tende ad assumere valori prossimi a quelli della prima classe d'età; se però nella fascia 0-14 la malattia che quasi unicamente colpisce le vie respiratorie

è l'asma, negli anziani cresce notevolmente il peso di patologie croniche più gravi, come la BPCO.

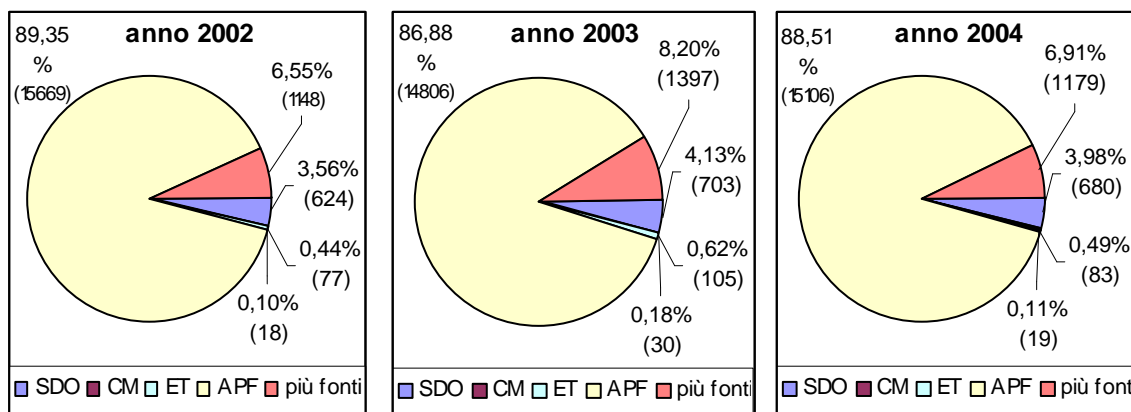
Nel periodo di tempo studiato, i tassi standardizzati presentano un andamento decrescente per entrambi i sessi, condizionati dalle classi di età più estreme, con diminuzione tra 2002 e 2003 significativa al 95%.

Tab. 3-MPCO: Numero assoluto e percentuale dei casi di MPCO identificati da fonte SDO, CM, APF, ET e più fonti (contributo esclusivo), per anno e classi d'età. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004

fonte	classi d'età											
	0-14		15-34		35-54		55-64		65-74		75+	
	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%
	2002											
SDO	16	0,41%	12	0,46%	29	0,94%	46	2,16%	139	5,15%	382	12,56%
CM	0	0,00%	0	0,00%	0	0,00%	0	0,00%	1	0,04%	17	0,56%
APF	3784	95,89%	2479	94,08%	2932	95,01%	1952	91,77%	2326	86,15%	2196	72,19%
ET	16	0,41%	18	0,68%	9	0,29%	10	0,47%	9	0,33%	15	0,49%
più fonti	130	3,29%	126	4,78%	116	3,76%	119	5,59%	225	8,33%	432	14,20%
tot	3946	100,00%	2635	100,00%	3086	100,00%	2127	100,00%	2700	100,00%	3042	100,00%
	2003											
SDO	19	0,49%	18	0,73%	23	0,77%	49	2,49%	139	5,42%	455	14,44%
CM	0	0,00%	1	0,04%	1	0,03%	2	0,10%	1	0,04%	25	0,79%
APF	3706	95,05%	2292	92,68%	2782	93,17%	1770	89,89%	2104	82,09%	2152	68,30%
ET	20	0,51%	24	0,97%	21	0,70%	5	0,25%	18	0,70%	17	0,54%
più fonti	154	3,95%	138	5,58%	159	5,32%	143	7,26%	301	11,74%	502	15,93%
tot	3899	100,00%	2473	100,00%	2986	100,00%	1969	100,00%	2563	100,00%	3151	100,00%

		2004											
SDO	14	0,37%	15	0,61%	24	0,78%	46	2,40%	152	5,88%	429	13,18%	
CM	0	0,00%	0	0,00%	2	0,07%	0	0,00%	3	0,12%	14	0,43%	
APF	3589	95,07%	2295	93,07%	2897	94,43%	1743	90,92%	2212	85,50%	2370	72,83%	
ET	11	0,29%	23	0,93%	14	0,46%	9	0,47%	6	0,23%	20	0,61%	
più fonti	161	4,26%	133	5,39%	131	4,27%	119	6,21%	214	8,27%	421	12,94%	
tot	3775	100,00%	2466	100,00%	3068	100,00%	1917	100,00%	2587	100,00%	3254	100,00%	

Grafico 1-MPCO: Distribuzione percentuale del contributo esclusivo delle fonti SDO, CM, APF, ET e di più fonti, per anno di stima. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.



Il contributo delle fonti per anni di stima e classi d'età mostra come, anche in questo caso, i farmaci siano decisivi per l'identificazione dei soggetti; diversamente dall'asma, però, il contributo dell'APF diminuisce al crescere dell'età, essendo sostituito dalle SDO e da informazioni ricavate da più fonti archivi.

Non si nota una tendenza nello scarso contributo esclusivo delle esenzioni ticket per sesso, classi d'età o anni di stima; complessivamente, le ET contribuiscono alla cattura del 3% circa dei soggetti.

Le morti per malattie polmonari non trattate nel corso del periodo immediatamente precedente sembrano sostanzialmente nulle, tranne che negli anziani: probabilmente, suddividendo ulteriormente la classe d'età

75+, si potrebbe apprezzare un andamento più graduale del fenomeno, come emerso nel caso della BPCO. Analogamente, la brusca crescita del contributo esclusivo delle SDO è amplificata dall'ampiezza dell'ultima classe di età, per la quale aumenta anche la quota di casi individuati da più fonti, probabilmente a causa della cronicità della malattia polmonare, che tende a progredire.

Rispetto all'algoritmo utilizzato per l'asma, nel caso delle MPCO sono state aggiunte altre categorie di farmaci; il contributo specifico per tipologia è riportato nella tabella seguente:

Tab. 4-MPCO: Contributo esclusivo per tipologia di farmaco, per anno. La percentuale è calcolata sul contributo assoluto della fonte APF. Periodo 2002-2004. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004

anno	codici ATC												totale	
	R03A*		R03C*		R03DC*		R03BB*		R03DA*		più farmaci			
	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%
2002	13667	81,44%	562	3,35%	0	0,00%	426	2,54%	0	0,00%	2126	12,67%	16781	100,00%
2003	12949	80,13%	287	1,78%	0	0,00%	531	3,29%	0	0,00%	2393	14,81%	16160	100,00%
2004	13315	81,94%	85	0,52%	0	0,00%	589	3,62%	0	0,00%	2260	13,91%	16249	100,00%
totale	39931	81,18%	934	1,90%	0	0,00%	1546	3,14%	0	0,00%	6779	13,78%	49190	100,00%

Gli adrenergici per aerosol, farmaci perlopiù leggeri, prevalgono come nel caso dell'asma, essendo prevalentemente prescritti ai bambini (per i quali rappresentano oltre il 90% della terapia farmacologica); l'interazione di più farmaci, che rassicura sull'effettiva pertinenza dei casi considerati prevalenti, costituisce in media quasi il 14% dei casi; dall'osservazione dell'andamento per età, non riportato, si nota come la tendenza ad assumere più farmaci specifici cresca con l'età, subentrando all'asma forme più gravi. Infine, il contributo esclusivo per codice di diagnosi alla dimissione da strutture ospedaliere, riportato in tabella 5, rivela come in quasi 4 casi su 5 i

codici utilizzati siano l'ICD IX 491 e 492 corrispondenti rispettivamente a bronchite cronica ed enfisema.

Un aspetto positivo dal punto di vista della qualità dei dati è la percentuale di codici generici (ICD IX 490, 494, 496), che presenta il contributo più basso fra i codici singoli e un trend temporale decrescente.

Tab. 5-MPCO: Contributo esclusivo per codice di dimissione, per anno. La percentuale è calcolata sul contributo assoluto della fonte SDO. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

anno	codici ICD IX									
	491-492		490, 494, 496		493		più codici		totale	
	n°	%	n°	%	n°	%	n°	%	n°	%
2002	1088	75.35%	134	9.28%	145	10.04%	77	5.33%	1444	100.00%
2003	1200	78.64%	96	6.29%	152	9.96%	78	5.11%	1526	100.00%
2004	1129	80.53%	61	4.35%	144	10.27%	68	4.85%	1402	100.00%
totale	3417	78.16%	291	6.66%	441	10.09%	223	5.10%	4372	100.00%

La distribuzione dei codici ICD IX per età, non riportata, conferma quanto atteso: asma e BPCO mostrano tendenze opposte rispetto all'età.

Confrontando i dati fra i centri si sono riscontrati valori simili dei tassi standardizzati, ma per Venezia l'andamento per età non è conforme alle altre aree, presentando tassi specifici molto più elevati nei giovani e più bassi negli anziani.

La sovrapposizione delle diagnosi di asma e BPCO è rilevante nelle classi d'età centrali, con picco fra 35 e 54 anni, fascia in cui è maggiormente difficile distinguere le due patologie in un soggetto.

I risultati sono complessivamente in linea con quanto atteso.

Diabete

Il diabete è una patologia piuttosto diffusa nella popolazione, che può con il tempo provocare gravi impedimenti alla vita del malato; l'Organizzazione Mondiale della Sanità ne raccomanda il monitoraggio e la prevenzione, trattandosi di una malattia in crescita lenta ma regolare nella popolazione del mondo occidentale.

Il gruppo di lavoro ha definito il seguente algoritmo per selezionare i diabetici prevalenti:

FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
APF	A10		almeno 2 prescrizioni nell'anno di stima
SDO	250 ²³	tutte	soggetti dimessi nell'anno di stima e nei 4 anni precedenti, non deceduti (per tutte le cause) al 1 gennaio dell'anno di stima
ET	013.250		soggetti esenti nell'anno di stima e nei 3 anni precedenti, non deceduti (per tutte le cause) al 1 gennaio dell'anno di stima

Viene dichiarato prevalente un individuo che nell'anno di stima presenti almeno una delle 3 condizioni sopra elencate. Quest'ultimo aggiornamento del protocollo non include l'archivio dei CM nell'algoritmo, poichè si ritiene inverosimile che al giorno d'oggi un soggetto muoia di diabete senza che questa malattia sia stata precedentemente diagnosticata e abbia dunque lasciato traccia negli altri archivi; in ogni caso, i malati gravi di diabete sono sottoposti a terapia farmacologica, quindi rintracciabili dall'APF anche se non esenti o ricoverati nel periodo considerato.

La cattura di casi soltanto tramite fonte CM potrebbe essere indicatore di scarsa qualità degli archivi; comunque, quando nel precedente protocollo i

²³ ICD IX 250: Diabete mellito.

certificati di morte erano inclusi tra le fonti, il loro contributo esclusivo era praticamente nullo (inferiore allo 0,1%).

Si è deciso di considerare il contributo longitudinale delle fonti SDO ed ET per catturare anche i casi stabilizzati di diabete che potrebbero non figurare se ci si limitasse ai record relativi all'anno di stima: in questo modo non verrebbero considerati i diabetici già esenti, catturando solamente le nuove prescrizioni nell'anno, e chi in passato avesse avuto una fase acuta della malattia, ricorrendo al ricovero, ma trattandola ora semplicemente con terapia dietetica.

La fonte SDO contribuisce per i 4 anni precedenti quello di stima, mentre la fonte ET per 3 anni complessivamente. La scelta del periodo è stata dettata dalla disponibilità di archivi di buona qualità presso i centri partecipanti, volendo conformare l'ampiezza della finestra temporale per ragioni di confrontabilità.

Utilizzando i dati archiviati a scopo amministrativo si perdono sicuramente i diabetici lievi trattati solamente mediante una dieta appropriata; inoltre, non sempre è conveniente dichiarare la malattia, pur potendo ottenere un sostegno economico per la sua cura. Tra le disfunzioni causate dal diabete, infatti, vi è una perdita di controllo del sensoriale e per questo motivo chi soffre di diabete non è abilitato alla guida dei veicoli. Per non perdere la patente, chi può basare la sua cura sulla dieta controllata ricorrendo saltuariamente a farmaci preferisce spesso non chiedere l'esenzione dal ticket.

Tab. 1-diabete: Riepilogo dei casi prevalenti di diabete, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

sesso	anno	classi d'età						totale
		0-14	15-34	35-54	55-64	65-74	75+	
uomini	2002	15	127	1.195	2193	2502	1787	7819
	2003	13	136	1.166	2202	2603	1870	7990
	2004	20	146	1.105	2219	2644	1974	8108
tot		48	409	3466	6614	7749	5631	23917
donne	2002	11	94	558	1416	2214	2750	7043
	2003	10	99	552	1423	2193	2887	7164
	2004	13	118	555	1396	2247	2974	7303
tot		34	311	1665	4235	6654	8611	21510
totale	2002	26	221	1.753	3609	4716	4537	14862
	2003	23	235	1.718	3625	4796	4757	15154
	2004	33	264	1660	3615	4891	4948	15411
tot		82	720	5131	10849	14403	14242	45427

Tab. 2-diabete: Tassi di prevalenza (valori *100) stimati per diabete, per sesso, classi di età e anno di prevalenza. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004. Standardizzazione con la popolazione italiana al censimento 2001, fonte ISTAT.

sesso	anno	classi d'età						Tasso Grezzo pop al 30/06	Tasso Stand.	I.C. 95%
		0-14	15-34	35-54	55-64	65-74	75+			
uomini	2002	0,09	0,36	2,74	10,44	15,66	16,12	5,45	5,09	4,99 - 5,20
	2003	0,08	0,40	2,65	10,45	16,06	16,30	5,55	5,15	5,04 - 5,25
	2004	0,11	0,44	2,48	10,58	16,03	16,69	5,60	5,16	5,06 - 5,27
tot								5,53	5,13	

donne	2002	0,07	0,29	1,27	6,03	10,59	12,69	4,45	3,31	<i>3,23 - 3,39</i>
	2003	0,06	0,32	1,24	6,08	10,42	13,01	4,52	3,33	<i>3,26 - 3,41</i>
	2004	0,08	0,38	1,23	6,04	10,62	13,15	4,58	3,38	<i>3,30 - 3,45</i>
tot								4,52	3,34	
totale	2002	0,08	0,33	2,00	8,11	12,78	13,85	4,93	4,11	<i>4,05 - 4,18</i>
	2003	0,07	0,36	1,94	8,15	12,87	14,13	5,01	4,15	<i>4,09 - 4,21</i>
	2004	0,10	0,41	1,85	8,20	12,99	14,36	5,07	4,18	<i>4,12 - 4,24</i>
tot								5,00	4,14	

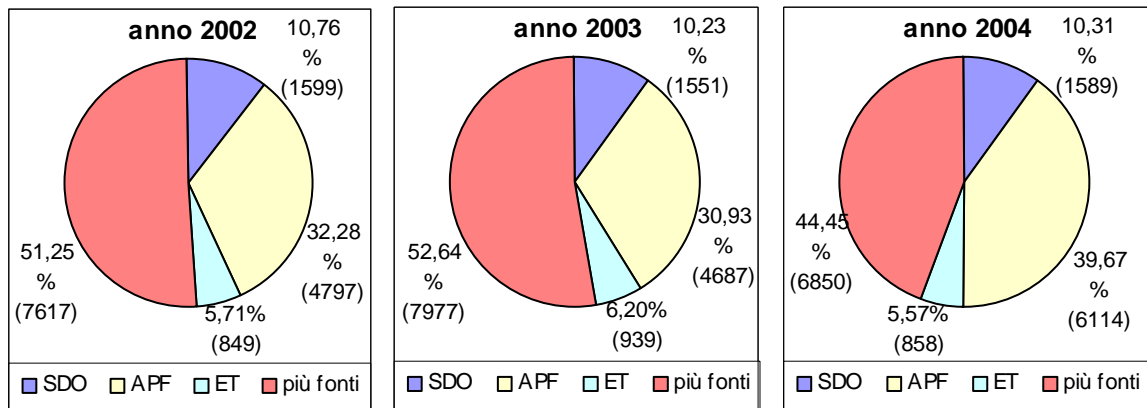
Osservando l'andamento dei tassi di prevalenza riportati nella tabella 2 si nota una maggior frequenza della malattia negli uomini rispetto alle donne e un trend crescente con l'avanzare dell'età. I tassi di prevalenza rimangono più bassi per le donne rispecchiando l'andamento di quelli maschili, tranne che nella classe 75+ anni, che riporta un incremento notevole nel tasso specifico femminile, ma limitato negli uomini.

Nei 3 anni di stima, la porzione di malati nella popolazione tende ad aumentare (tra 2002 e 2004 +1,6% nei uomini, +1,98% nelle donne), seppur con differenze non significative nel periodo. Questa tendenza si riscontra principalmente nella fascia d'età più anziana (oltre 3,5% di incremento tra 2002 e 2004) e soprattutto tra 15 e 34 anni (tassi del 2004 più elevati persino del 21% negli uomini e del 32% nelle donne in confronto a quelli di due anni prima), con consistenza simile fra uomini e donne. Analizzando il contributo esclusivo degli archivi che concorrono alla definizione dei casi si prova ad esplorare la motivazione di questo aumento, tentando di capire se la stima sia influenzata dalle fonti utilizzate o sia invece attendibile.

Tab. 3-diabete: Numero assoluto e percentuale dei casi di diabete identificati da fonte SDO, APF, ET e più fonti (contributo esclusivo), per anno e classi d'età. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.

fonte	classi d'età											
	0-14		15-34		35-54		55-64		65-74		75+	
	n°	%	n°	%	n°	%	n°	%	n°	%	n°	%
2002												
SDO	1	3,85%	28	12,67%	126	7,19%	271	7,51%	376	7,97%	797	17,57%
APF	3	11,54%	60	27,15%	560	31,95%	1217	33,72%	1636	34,69%	1321	29,12%
ET	1	3,85%	26	11,76%	202	11,52%	308	8,53%	180	3,82%	132	2,91%
più fonti	21	80,77%	107	48,42%	865	49,34%	1813	50,24%	2524	53,52%	2287	50,41%
tot	26	100,00%	221	100,00%	1753	100,00%	3609	100,00%	4716	100,00%	4537	100,00%
2003												
SDO	0	0,00%	18	7,66%	129	7,51%	220	6,07%	386	8,05%	798	16,78%
APF	1	4,35%	63	26,81%	561	32,65%	1195	32,97%	1550	32,32%	1317	27,69%
ET	0	0,00%	27	11,49%	212	12,34%	317	8,74%	227	4,73%	156	3,28%
più fonti	22	95,65%	127	54,04%	816	47,50%	1893	52,22%	2633	54,90%	2486	52,26%
tot	23	100,00%	235	100,00%	1718	100,00%	3625	100,00%	4796	100,00%	4757	100,00%
2004												
SDO	0	0,00%	20	7,58%	117	7,05%	214	5,92%	395	8,08%	843	17,04%
APF	12	36,36%	97	36,74%	633	38,13%	1545	42,74%	2071	42,34%	1756	35,49%
ET	1	3,03%	46	17,42%	225	13,55%	282	7,80%	208	4,25%	96	1,94%
più fonti	20	60,61%	101	38,26%	685	41,27%	1574	43,54%	2217	45,33%	2253	45,53%
tot	33	100,00%	264	100,00%	1660	100,00%	3615	100,00%	4891	100,00%	4948	100,00%

Grafico 1-diabete: Distribuzione percentuale del contributo esclusivo delle fonti SDO, APF, ET e di più fonti, per anno di stima. Popolazione dell'Azienda ULSS 12 veneziana, 2002-2004.



Dal grafico, si nota subito la forte diminuzione del contributo esclusivo di più fonti nel 2004 rispetto ai due anni precedenti, compensato dall'aumento di quasi 9 punti percentuali del contributo dei farmaci dal 2003, che si riscontra in tutte le fasce d'età e in entrambi i sessi (cfr. tabella 3). Nell'arco di un anno il numero assoluto dei farmaci antidiabetici non associati ad altre cure comprese nell'algoritmo aumenta del 30%; questo dato è di difficile interpretazione, poiché potrebbe essere indicatore di una maggior frequenza della malattia in una forma non leggera nella popolazione oppure di un problema di qualità degli archivi.

In realtà, considerando il contributo assoluto delle 3 fonti, si osserva nel 2004 una diminuzione di casi alla cui identificazione contribuisce l'archivio delle ET, mentre SDO e APF non subiscono variazioni di rilievo. Questo spiega verosimilmente quanto osservato sopra: diminuisce il numero di casi identificati anche da esenzione (dal 39% del 2003 a un quarto nel 2004) ma i diabetici aumentano anziché diminuire, emergendo dalle prescrizioni farmaceutiche che presentano un contributo assoluto e soprattutto esclusivo più alto.

Negli ultrasessantacinquenni il diabete viene individuato esclusivamente tramite SDO in media nel 17% dei casi, percentuale più alta che in tutte le

altre età e più che raddoppiata rispetto alla classe immediatamente precedente; questo dato risente sicuramente della peculiarità delle esenzioni ticket, spettanti di diritto oltre i 65 anni.

Si era osservato nel periodo un incremento notevole della prevalenza nella fascia 15-34 anni; per questa classe il contributo delle diverse fonti si modifica soprattutto nel 2004, aumentando decisamente la percentuale di diabetici identificati con i farmaci e con le ET e diminuendo di molto il contributo multiplo e delle SDO.

I diabetici oltre i 65 anni identificati esclusivamente come nuovi esenti sono molto pochi e questo dato è conforme a quanto atteso, subentrando l'esenzione per età.

Queste dinamiche suggeriscono che l'algoritmo e gli archivi, in particolar modo ET e APF, siano ancora in una fase di assestamento.

Il trend della prevalenza di diabete è comparabile nei vari centri e in generale non riscontrato in dati forniti da altre fonti; ciò porta a pensare che l'algoritmo possa introdurre una sovrastima dei casi, pur limitandosi a catturare i diabetici gravi, ovvero quelli che necessitano di cure sanitarie e non possono limitarsi a seguire diete specifiche. La stima nel 2004 è più in linea con dati provenienti da altri studi, e per Venezia il tasso standardizzato è sovrapponibile a quello fornito dall'ISTAT relativamente al 2005²⁴; questa convergenza non può che essere di buon auspicio sulla qualità delle stime prodotte applicando l'algoritmo, nonostante rimangano aspetti da meglio chiarire.

²⁴ Cfr. Bibliografia: ISTAT, 2005

Tumore della mammella e tumore del colon-retto

Le potenzialità dello sfruttamento degli archivi elettronici per la stima di incidenza e prevalenza delle malattie sono state inizialmente esplorate nell'ambito dei registri tumori per la registrazione automatica dei casi (cfr. pag. 26).

Considerata l'importanza di monitorare la diffusione delle patologie tumorali nella popolazione, il gruppo di lavoro nazionale AIE-SISMEC ha ritenuto di includere due tipi di tumore fra le patologie scelte per la sperimentazione della metodologia, e precisamente il tumore della mammella, particolarmente presente nella popolazione femminile, e quello del colon-retto, che colpisce entrambi i sessi, con frequenza maggiore nei maschi.

L'avanzamento del lavoro per queste due patologie neoplastiche è relativo al gennaio 2007, momento in cui si è deciso di accantonare momentaneamente le stime per i tumori in seguito alle numerose problematiche emerse, che richiederanno un'attenzione e un impegno più approfonditi.

In questa sede verrà presentato l'ultimo aggiornamento dell'algoritmo, non definitivo, e ci si soffermerà brevemente sulla discussione delle problematiche tuttora rimaste irrisolte; non verranno presentate le stime perché scarsamente attendibili.

L'obiettivo del gruppo era fornire stime di incidenza annuale per le due patologie; le diverse esperienze maturate dai partecipanti hanno portato alla definizione di più algoritmi per ciascun tumore, al fine di valutare quale potesse essere il più adeguato.

Viene in questa sede utilizzato anche l'archivio dei referti di anatomia patologica (AP), che contiene informazioni determinanti nell'identificazione dei malati.

Nel caso del cancro della mammella, sono state proposte le seguenti 2 combinazioni di fonti:

ALGORITMO	FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
1	AP	174		con morfologia /3 o /2
	MP	174, 233	principale	con associata procedura chirurgica 85.12, 85.20-25, 85.33-36, 85.41-48, esclusi i soggetti con ICD IX V10.3 nelle diagnosi secondarie
2	SDO	174, 233	principale	con associata procedura chirurgica 85.12, 85.20-25, 85.33-36, 85.41-48, esclusi i soggetti con ICD IX V10.3 nelle diagnosi secondarie

La scelta di utilizzare due differenti algoritmi deriva dalla diversa disponibilità di dati di buona qualità nei vari centri. Per quanto riguarda la SDO, laddove la compilazione del campo relativo ad intervento chirurgico non sia affidabile, si utilizza l'algoritmo (1): in questo modo, essendo la maggior parte dei casi (circa 98%) rintracciabile dall'AP, viene individuata la maggior parte di soggetti ricoverati in strutture dell'Azienda Sanitaria di riferimento, utilizzando l'archivio della MP per recuperare coloro che sono stati assistiti presso un'altra ULSS.

Poiché l'archivio dei referti di AP non è disponibile ovunque, si utilizza l'algoritmo (2) nel caso in cui l'affidabilità della SDO sia buona.

L'evento incidente è definito, a seconda degli algoritmi, come:

- (1) prima diagnosi istopatologica con codice 174/3 o 174/2 MP;
- (2) primo ricovero.

L'intervallo di tempo per ripulire l'incidenza dalla prevalenza nel caso di due eventi attribuiti allo stesso soggetto è di 60 mesi, con controllo su:

- (1) precedenti referti di AP, SDO per i codici specifici in qualsiasi posizione di diagnosi, anche per interventi in situ;
- (2) SDO per i codici specifici in qualsiasi posizione di diagnosi, anche per interventi in situ.

Una problematica generale che rimane irrisolta riguarda l'ampiezza del periodo per ripulire i casi incidenti dai prevalenti, poiché è stata avanzata l'osservazione che non siano sufficienti 60 mesi per definire due casi come distinti: la stessa malattia può ripresentarsi anche a distanza di molto tempo dopo l'intervento chirurgico, essendo caratterizzata da possibili ricadute.

Utilizzando la prima procedura, nell'AULSS veneziana oltre il 60% dei casi viene individuato soltanto dall'archivio di AP, e la percentuale aumenta con l'età dei soggetti; soltanto nel 17% dei casi le informazioni provengono da entrambe le fonti.

Inoltre, si è notato come le due procedure portino a risultati diversi: nel caso di Venezia, nell'anno 2003 sono stati individuati complessivamente 377 soggetti, dei quali 284 con entrambi i metodi, mentre 74 solo dall'archivio di AP (algoritmo 1) e 19 identificati solo dall'archivio SDO (algoritmo 2).

Le questioni che rimangono insolte sono molteplici:

- Come considerare i soggetti che compaiono soltanto con un metodo?
- E' opportuno scegliere uno dei due algoritmi, escludendo dunque sistematicamente una parte di soggetti, ed eventualmente quale dei due?
- E' possibile quantificare l'errore di stima introdotto dai due metodi?
- Come stabilire quale dei due metodi ha un potere discriminante maggiore?
- Si pensa di stabilire una nuova combinazione di fonti che tenga conto delle due sperimentate?

Gli algoritmi utilizzati per la stima dei casi incidenti di cancro al colon-retto sono i seguenti:

ALGORITMO	FONTE	CODICI	DIAGNOSI	PARTICOLARITA'
1	AP	153, 154		con morfologia /3
	MP	153, 154	principale (DRG chirurgico)	esclusione soggetti con ICD IX V10.05-06 nelle diagnosi secondarie
2	SDO	153, 154	principale (DRG chirurgico)	esclusione soggetti con ICD IX V10.05-06 nelle diagnosi secondarie

Le ragioni della definizione di due algoritmi diversi sono le stesse presentate per il cancro della mammella (cfr. pag. 83).

L'evento incidente è definito, a seconda del metodo adottato, come:

- (1) prima diagnosi istopatologica con codice 153-154/3 o MP;
- (2) primo ricovero.

L'intervallo di tempo per ripulire l'incidenza dalla prevalenza nel caso di due eventi attribuiti allo stesso soggetto è di 60 mesi, con controllo su:

- (1) precedenti referti di AP, SDO per i codici specifici in qualsiasi posizione di diagnosi, anche per interventi in situ;
- (2) SDO per i codici specifici in qualsiasi posizione di diagnosi, anche per interventi in situ.

Sono emersi problemi riguardanti i malati non trattati, specialmente in età avanzate, che non possono dunque essere identificati mediante questi algoritmi; ulteriori complicazioni si presentano nel caso della concomitanza di tumori anche in altre sedi e quindi difficilmente individuabili come cancro

del colon-retto, e di codifiche poco attendibili dell'AP, specialmente per alcuni centri.

Dal confronto fra i dati del centro di Torino ottenuti con il secondo algoritmo e quelli forniti dal Registro Tumori della medesima area (considerato gold standard) emerge una sottostima sistematica della fonte SDO negli anziani; inoltre, del 40% dei casi considerati incidenti dal Registro ma non individuati dal centro perché non soddisfacenti i criteri dell'algoritmo, il 24% presentava DRG medico anziché chirurgico.

Anche in questo caso, è più opportuno scegliere uno dei due algoritmi, ed eventualmente quale dei due, oppure stabilire una nuova combinazione di fonti che tenga conto delle due sperimentate e delle criticità emerse?

In entrambi i casi sono state riscontrate difficoltà nel confrontare i risultati ottenuti dai vari centri, perché si è fatto riferimento a periodi disomogenei e gli archivi di AP avevano diversi livelli di completezza; inoltre, i dati forniti dai Registri Tumori esistenti erano standardizzati con popolazioni diverse, e dunque anch'essi difficilmente confrontabili.

Come già anticipato, gli algoritmi dovranno essere ampiamente perfezionati.

Aspetti da approfondire

Il gruppo di lavoro ha prodotto algoritmi che si sono dimostrati complessivamente di buon funzionamento per il raggiungimento degli obiettivi prefissati.

Permangono notevoli difficoltà interpretative, alcune peculiari di ogni patologia, come discusso nelle apposite sezioni, altre di carattere più generale.

Non è chiaro se differenze riscontrate fra i centri siano dovute a diversi livelli di qualità degli archivi, oppure catturino un'effettiva variabilità spaziale della frequenza delle malattie, o ancora riflettano attitudini terapeutiche non omogenee, ipotesi peraltro assai plausibili. Tutti questi fattori sono verosimilmente concomitanti; nel futuro si dovrà cercare di approfondire queste dinamiche per intervenire dove sia possibile allo scopo di affinare i metodi impiegati, mantenendo il requisito fondamentale di confrontabilità e controllabilità dei dati prodotti.

Per quanto riguarda Venezia, va approfondito il meccanismo riguardante le esenzioni ticket, poiché rispetto agli altri centri questa fonte contribuisce in modo costantemente maggiore nell'identificazione dei soggetti a tutte le età. Risultano in particolar modo poco comprensibili le richieste di esenzione per malattia attribuite agli ultrasessantacinquenni, spettando di diritto per età.

Sarà importante valutare le caratteristiche dei dati contenuti in quest'archivio nel lungo periodo di disponibilità, che ha inizio nel 1991. Emergeranno tendenze particolari di ricorso all'esenzione nell'AULSS veneziana? Si tratta piuttosto di anomalie di archiviazione e gestione dei dati, come ad esempio scarso aggiornamento, che portano ad una sovrastima degli esenti in vita? Nel confronto con gli altri centri, non sembra però che un'eventuale singolarità abbia effetti di distorsione delle stime prodotte. Questa considerazione mette in luce un fondamentale punto di forza di questo metodo di cattura dei casi: l'integrazione di molteplici fonti indipendenti, che permette di intercettare al meglio i malati seppur in condizioni diagnostico-terapeutiche mutevoli. Come si è notato dall'analisi dei contributi degli archivi, diversità si riscontrano non soltanto a livello spaziale, ma anche temporale in una medesima realtà: così, gli operatori che vogliono dedicarsi alla stima della frequenza delle malattie nel tempo in una popolazione pur ristretta ne devono tener conto, valutando la qualità degli archivi e le dinamiche, anche di carattere burocratico-amministrativo, che

possono influenzare la visibilità dei casi attraverso alcune basi di dati piuttosto che altre.

Un'ulteriore questione che rimane in sospeso riguarda le implicazioni dell'utilizzo longitudinale delle fonti SDO ed ET: coloro che risultavano residenti presso l'Azienda ULSS al momento del contatto con il servizio sanitario sono ancora residenti presso la stessa, potendo essere inclusi nel calcolo della prevalenza, oppure sono emigrati senza che sia rimasta traccia dello spostamento? Si tratta del problema dell'aggiornamento dell'Anagrafe Sanitaria, che spesso riporta date di inizio e fine assistenza poco attendibili. L'ASL di Torino sta ora provando a stimare il peso, in assenza del controllo sulla residenza, della sovrastima della casistica da emigrati non esclusi e della sottostima da immigrati non inclusi in Anagrafe, ma la verifica è difficilmente effettuabile dai singoli centri per ragioni di disponibilità dei dati e del tempo richiesto; sarà dunque da rimandare ad un prossimo futuro perfezionamento della metodologia proposta.

Terminata la panoramica sui risultati del lavoro di questo gruppo nazionale, si tratta ora di valutare come si inserisce l'esperienza condotta nello scenario epidemiologico attuale e che valenza può avere in un'ottica di sviluppo della sanità pubblica italiana. Nel prossimo capitolo si proverà ad affrontare la questione, con un approccio che non ha la pretesa di essere esaustivo, ma semplicemente intende offrire degli spunti per ulteriori approfondimenti in questo ampio contesto emergente.

Capitolo 4

DISCUSSIONE

Come già accennato nell'introduzione, la possibilità di utilizzare le basi di dati elettroniche di popolazione per scopi epidemiologici si è profilata negli ultimi decenni, grazie principalmente all'accrescimento della potenza dei supporti tecnologici, che ha permesso di archiviare un'enorme quantità di informazioni senza per questo dover disporre di computer di particolare potenza.

In sanità pubblica si sono sviluppati diversi approcci per lo sfruttamento di questa possibilità, e tendenze differenti sono state intraprese nei vari contesti nazionali ed internazionali.

Di seguito, si riporteranno a titolo esemplificativo e assolutamente non esaustivo alcune situazioni, rappresentative delle principali tendenze in atto in ambito italiano ed europeo.

Registri Tumori

L'ambito dei Registri Tumori è stato e rappresenta tuttora un laboratorio per l'esplorazione delle potenzialità dell'archiviazione elettronica, ponendo le basi per l'estensione delle procedure ad altre patologie.

In Italia soltanto il 28% della popolazione è coperta da Registri, generali o specifici, per la stima dell'incidenza delle neoplasie, con grandi disparità geografiche, essendo maggiormente diffusi al centro e al nord, dove sono estesi anche ad intere Regioni (si veda l'esempio del Friuli - Venezia Giulia), e scarsamente nel sud del Paese.

Il problema di fondo di questi centri è il loro carattere di spontaneità: la legge italiana non prevede, infatti, l'obbligo di monitoraggio delle patologie tumorali per mezzo di questo strumento e la sua implementazione è lasciata alla sensibilità di singoli medici o equipe.

La disponibilità di flussi informativi sempre più completi e integrati ha reso possibile la creazione di sistemi per fornire un supporto di registrazione automatica alla fase di identificazione dei casi; questo modello viene applicato attualmente in 7 Registri Tumori generali e permette di alleggerire il carico di lavoro manuale in tutti i casi di semplice identificazione del soggetto malato (cfr. pag. 26).

In ambito europeo la situazione è diversificata, sia dal punto di vista dei metodi di registrazione utilizzati che per la diffusione geografica sui territori nazionali.

Se in alcuni Paesi, quali Francia, Spagna e Portogallo, i Registri Tumori si sono sviluppati in modo non omogeneo sul territorio, similmente a quanto visto per la contesto italiano, vi sono popolazioni interamente coperte da Registri. E' il caso dei Paesi Scandinavi, che si sono dotati di Registri di popolazione già da decenni, ma anche del Regno Unito e dell'Irlanda, dove la registrazione automatica si sta progressivamente diffondendo, seppur con punti di partenza differenti.

La copertura totale della popolazione inglese risale già ad alcuni anni fa e la registrazione era stata necessariamente implementata con criteri manuali di identificazione dei casi; nella fase attuale, compresa l'importanza strategica per finalità epidemiologiche e di sanità pubblica dell'utilizzo dei metodi più avanzati per il monitoraggio della diffusione dei tumori, si sta procedendo a convertire l'intero sistema di registrazione da manuale ad automatico.

In Irlanda, invece, l'iniziativa dei Registri è recente e si è scelto di strutturare il sistema fin da subito come registrazione automatica.

Altre iniziative

Quantità notevolissime di dati riguardanti fenomeni estremamente diversificati sono rese disponibili, soprattutto per le popolazioni del mondo occidentale, da organismi nazionali e internazionali accreditati, enti certificati, istituti di ricerca, ecc.

Si distinguono raccolte occasionali, limitate ad un periodo contingente, e raccolte periodiche, che mettono a disposizione informazioni su uno stesso fenomeno in modo sistematico nel corso del tempo. Ancora, i dati possono essere disponibili ad un dettaglio territoriale elevato, come a livello comunale o raramente sub-comunale (sezione di censimento), oppure essere più articolati ma riferiti a macro-aree di ampiezza almeno provinciale.

La maggior parte dei dati raccolti e messi a disposizione da queste strutture è di carattere aggregato, ovvero l'unità di analisi è la popolazione generale o un suo sottoinsieme: essi esprimono la diffusione di un fenomeno nel gruppo sotto forma di proporzione.

Di diversa portata informativa sono invece i dati individuali, che richiedono una complessità maggiore di raccolta ma permettono di valutare non soltanto la compresenza di fenomeni nella popolazione, bensì anche eventuali relazioni fra le variabili.

Ad esempio, confrontando due popolazioni, si può affermare che nell'insieme A, sul quale si è osservato un livello medio di istruzione più elevato, vi sia anche una proporzione maggiore di occupati rispetto all'insieme B; diversamente, si può sostenere che i soggetti con un titolo di studio più elevato siano più frequentemente occupati rispetto a chi possiede un titolo inferiore, e magari che la relazione sia proporzionale.

Nel primo caso si è in possesso di dati aggregati e ci si limita ad osservare la concomitanza di due caratteristiche *nella popolazione*: nulla permette di affermare che quanto osservato a livello generale sia effettivamente

riscontrabile a livello dei soggetti; nel secondo caso, disporre di dati per ogni *singolo individuo* rende possibile la conduzione di analisi più complesse con un apporto informativo più rilevante.

I metodi elaborati dal gruppo e fin qui presentati si basano sulla selezione dall'insieme della popolazione dei soli soggetti di interesse, sulla base di caratteristiche personali ricavate dagli archivi; poiché ad ogni individuo è assegnato un codice identificativo univoco, è assolutamente possibile condurre l'analisi sfruttando la portata informativa propria di dati individuali; allo stesso tempo, essendo il codice generato indipendentemente da informazioni anagrafiche personali, l'elaborazione rispetta la clausola dell'anonimato del soggetto.

Quest'ultima considerazione è cruciale per valutare l'avanzamento dell'utilizzo degli archivi elettronici sanitari per scopi epidemiologici in Europa. La situazione è infatti assai diversificata.

In alcuni Paesi, specialmente quelli Scandinavi, ai cittadini viene assegnato alla nascita un codice identificativo univoco con il quale sono registrate tutte le informazioni, di carattere economico, anagrafico, sanitario ecc., derivanti da contatti con servizi erogati dallo Stato. Questi archivi sono stati implementati già negli anni Cinquanta e permettono di avere ora a disposizione una ricca base di dati di elevata qualità e riferita ad un lungo periodo di tempo: le possibilità di condurre studi sulla popolazione o suoi sottogruppi sono praticamente illimitate. Il problema del rispetto della privacy essenzialmente non si presenta, poiché la questione è stata affrontata mediante l'assegnazione di codici anonimi generati casualmente, che permettono di mantenere contemporaneamente l'individualità dei dati e la riservatezza dei soggetti. La sicurezza resa possibile dalla tecnologia supporta la ricerca epidemiologica, creando un binomio dall'efficacia esemplare.

Situazioni diametralmente opposte si presentano in altri Stati, come ad esempio in Germania: le restrizioni imposte al trattamento dei dati personali sono talmente limitanti che di fatto impediscono qualsiasi utilizzo degli archivi elettronici, paralizzando lo sviluppo di questo innovativo filone di ricerca.

Certamente la tutela della riservatezza individuale, in particolar modo relativa a tematiche sensibili, è di fondamentale importanza in un Stato che voglia garantire la libertà ai suoi cittadini; d'altro canto, però, essi sono portatori di diritti e fra questi la salute è certamente uno dei più rilevanti: per rispettare questo diritto, lo Stato si deve far carico di attuare quanto è nelle sue possibilità.

E' evidente come la ricerca scientifica nel campo epidemiologico sia un mezzo utilizzabile per migliorare la qualità della salute della popolazione e per questo un'attività di estremo interesse pubblico; inoltre, come già affermato nell'introduzione, si caratterizza come strumento fondamentale per supportare la programmazione sanitaria, rendendo più efficiente ed efficace l'assistenza: ne consegue come sia irrinunciabile il superamento dell'ostacolo che impedisce lo sfruttamento di informazioni così preziose in nome del rispetto della privacy.

Se al centro dell'interesse di uno Stato è posto il cittadino, un Paese moderno si deve prodigare per stabilire un equilibrio fra componenti fondamentali coesistenti, proficuo sia per l'individuo che per le componenti stesse.

In questo contesto, l'Italia sta ponendo le basi per la risoluzione della questione: si è tentato di stilare il regolamento per il trattamento dei dati personali sensibili²⁵ in collaborazione con gli epidemiologi, per favorire la realizzazione degli studi nel pieno rispetto della privacy.

²⁵ "Codice in materia di protezione dei dati personali", D.L. 196, 30 giugno 2003.

Certamente, se queste regole fossero immutabili nel tempo perderebbero ben presto la loro forza innovativa: l'aggiornamento deve essere costante e attento, per adeguarle all'avanzamento tecnologico e scientifico, permettendo alla ricerca di operare serenamente nel pieno rispetto della legalità.

Il momento che l'epidemiologia sta attraversando nel nostro Paese può essere visto come una svolta: i presupposti per un significativo avanzamento di questa disciplina sono in parte stati posti, come l'esistenza diffusa di archivi di popolazione e la possibilità di un loro utilizzo regolamentato, in parte si stanno sviluppando, come l'integrazione delle basi di dati e il raggiungimento di una soglia accettabile di qualità.

E' importante domandarsi se la direzione che si sta percorrendo sia la più adeguata. Molteplici iniziative stanno emergendo a livello centrale e locale, presso organismi ed enti di ricerca, Università, Aziende Sanitarie. Numerose patologie sono oggetto di monitoraggio con metodi anche molto differenti fra loro, segno dell'interesse fervente per quest'ambito, ma sintomo anche di un inadeguato coordinamento che può portare a disperdere tempo ed energie. Uno sguardo alla molteplicità delle iniziative nell'ottica dell'uniformità dei metodi, per migliorare le procedure di raccolta dei dati e rendere confrontabili i risultati, è probabilmente l'atteggiamento più lungimirante da seguire.

Limitando l'attenzione ai metodi presentati nel capitolo precedente, approfondimenti e correzioni come visto si rendono necessari, ma le possibilità di estensione agli organi periferici del territorio nazionale non sono poi così remote.

Gli algoritmi per la definizione della casistica potrebbero essere ulteriormente perfezionati, introducendo nuove fonti informative,

modificando le procedure attuali o integrandole con altri metodi di selezione dei casi.

Ad esempio, archivi un tempo non utilizzabili, come quello delle visite specialistiche, potrebbero fornire maggior accuratezza e specificità ai dati raccolti; il passaggio dal numero di prescrizioni farmaceutiche al dosaggio somministrato al malato permetterebbe di distinguere con maggior attendibilità i casi gravi di malattia da quelli più lievi, con implicazioni positive soprattutto nelle patologie per le quali il contributo dell'APF è predominante; applicare il metodo di "cattura e ricattura"²⁶, che rende possibile stimare con l'equazione di Chapman anche i casi di malattia presenti nella popolazione ma non osservabili mediante gli archivi utilizzati, fornirebbe stime di incidenza e prevalenza meno distorte.

Prima ancora di ampliare definizioni, arricchire procedure e cimentarsi con altre patologie, per accreditare maggiormente questo modello di stima è necessario però stabilire dei requisiti minimi di qualità degli archivi, misurabili tramite indicatori condivisi dalla comunità scientifica, perché le stime prodotte in realtà e con metodi diversi possano essere attendibili e confrontabili.

²⁶ Cfr. bibliografia: Gallo U. et al., 2005.

Il metodo della "cattura e ricattura" parte dal presupposto che non tutti i casi di malattia presenti in una popolazione possano essere selezionati mediante l'utilizzo di archivi elettronici, essendo ad esempio soggetti ai quali non sia ancora stata diagnosticata la malattia, oppure trattati con una terapia che non lasci traccia negli archivi, ecc.

Per stimare più realisticamente la prevalenza di una patologia è stato proposto di avvalersi dell'equazione di Chapman:

dati due archivi indipendenti relativi ad una popolazione chiusa in cui tutti i soggetti abbiano la medesima probabilità di essere individuati e siano contraddistinti da un codice univoco, indicato con N1 il numero di soggetti catturati soltanto dall'archivio 1, N2 quelli catturati solo dall'archivio 2 ed M i malati ritrovati ("ricatturati") in entrambi gli archivi, il numero N complessivo dei prevalenti nella popolazione sarà così definito:

$$N = \left\{ \left[\frac{(N1+1)(N2+1)}{M+1} \right] - 1 \right\}$$

Il metodo è estendibile anche ai casi di impiego di più di due fonti.

Questa metodologia è raccomandata anche dall'Organizzazione Mondiale della Sanità per la produzione di stime più attendibili.

La costituzione di questi indicatori deve essere affidata ad organismi istituzionali competenti, come pure l'iniziativa di esaminare la metodologia proposta e valutarne una possibile estendibilità a livello nazionale.

Il SEI di Venezia

Nell'ottica di questo quadro, brevemente tracciato tentando di mettere in luce le dinamiche concernenti l'utilizzo di archivi elettronici di popolazione, il SEI di Venezia si pone come strumento estremamente interessante per rispondere alle esigenze emergenti.

Innanzitutto, l'integrazione sistematica di molti archivi rende possibile la conduzione di studi diversificati nel metodo e nelle finalità, senza richiedere un lavoro preparatorio ogni qualvolta si progetti una nuova attività di ricerca.

Questo aspetto, di notevole rilevanza dal punto di vista organizzativo, è emerso chiaramente durante lo svolgimento del lavoro presentato nel capitolo 3, dimostrandosi il SEI di Venezia una struttura organizzata in modo funzionale al lavoro richiesto, pur non essendo stata pensata appositamente per questo scopo.

La banca dati si è rivelata estremamente flessibile: può essere infatti sottoposta alle più svariate interrogazioni senza dover modificare l'architettura di integrazione degli archivi che la compongono, e nell'insieme dei centri facenti parte del sottogruppo per la produzione delle stime di frequenza di alcune malattie, Venezia è stata l'unica realtà in grado di fornire le stime di incidenza e prevalenza per tutte le patologie proposte, portando contemporaneamente avanti, sempre mediante il SEI, altri studi sui suoi assistiti.

Il lavoro preparatorio sulle fonti informative è infatti già stato effettuato e l'integrazione sistematica degli archivi permette di disporre di una banca dati di partenza costantemente aggiornata e pronta per essere interrogata.

Un vantaggio che questa struttura apporta è la possibilità di avere regolarmente sotto controllo i diversi archivi che la compongono, monitorandone la qualità e migliorando il record linkage fra essi.

L'esigenza principale che ha portato alla costruzione del SEI è stata la necessità di disporre di un supporto affidabile per la programmazione sanitaria nell'ambito dell'Azienda ULSS veneziana.

Da questo punto di vista, il sistema ha dimostrato di essere un valido sostegno, in grado di fornire tempestivamente dati di qualità riguardanti le condizioni di salute della popolazione, fruibile dagli organismi preposti alla pianificazione degli interventi sanitari dell'Azienda. Il monitoraggio continuo sugli assistiti, con un'ampia copertura di malattie, permette di disporre dei dati necessari per intervenire tempestivamente e con maggior efficacia sui problemi, attraverso l'attuazione di programmi di assistenza e prevenzione supportati da informazioni pressoché in tempo reale.

Sempre nell'ottica della gestione pubblica del sistema sanitario, il SEI ha dimostrato le sue notevoli e promettenti potenzialità quale strumento di sorveglianza della diffusione delle malattie nella popolazione, lasciando intravedere per il futuro opportunità di sviluppo estremamente interessanti.

I punti critici di questo sistema richiedono, come visto, ancora notevole lavoro, ma il fatto che siano emersi non può che costituire un aspetto positivo, rappresentando un punto di partenza essenziale per il perfezionamento del SEI.

Le possibilità di ampliarne l'utilizzo, impiegandone al meglio i punti di forza, fanno del sistema uno strumento fondamentale come supporto alla ricerca epidemiologica e al governo della sanità pubblica veneziana.

Tra le prospettive più innovative, la realizzazione di studi sullo sviluppo delle malattie a livello di popolazione, indagando in prospettiva longitudinale gli eventi a carico di soggetti malati registrati nei vari archivi, per ricostruire i profili di malattia degli assistiti e analizzarli in chiave assistenziale ed eziologia.

Così, si può selezionare ad esempio la coorte degli infartuati ad una certa data, risalire nel tempo individuando gli altri eventi registrati per questi soggetti e provare ad affrontare questioni di grande interesse epidemiologico e sanitario, quali l'esistenza di circostanze ricorrenti fra i malati, l'individuazione di fattori che sembrano condizionare la probabilità di ammalarsi di infarto, il caratterizzarsi di questa patologia come fattore di rischio a sua volta predisponente ad altre malattie.

L'analisi potrà essere condotta su patologie specifiche, su fasce d'età (ad esempio, sulla popolazione anziana) o ancora sull'intera categoria dei malati, stimando la prevalenza di soggetti affetti da una qualche forma di infermità nella popolazione.

Allo stato attuale, gli anni di disponibilità di dati attendibili ricavabili dagli archivi sono limitati (cfr. pag. 12), ma nel futuro il SEI renderà possibile esaminare le dinamiche di salute degli assistiti sistematicamente organizzati in coorte, senza dover ricorrere alla progettazione di onerose indagini ad hoc, permettendo un'osservazione longitudinale e continua della popolazione.

Con il delinearsi di questo ulteriore utilizzo del SEI, l'AULSS veneziana disporrà di notevoli strumenti innovativi per ampliare la ricerca epidemiologica e supportare la gestione sanitaria.

Capitolo 5

CONCLUSIONI

L'obiettivo che si proponeva questo elaborato era valutare le potenzialità e i limiti del monitoraggio della frequenza delle malattie nella popolazione mediante un sistema epidemiologico integrato.

I metodi presentati nel capitolo 3 hanno mostrato la loro pertinenza per stimare incidenza e prevalenza di alcune patologie nella popolazione, raggiungendo nel complesso le finalità prefissate; sono emerse dinamiche da chiarire e aspetti da perfezionare, ma in generale i risultati ottenuti sono stati soddisfacenti.

Il SEI di Venezia si è dimostrato una struttura organizzata in modo funzionale al lavoro richiesto, pur non essendo stata pensata appositamente per questo scopo.

Il Sistema Epidemiologico Integrato ha mostrato di essere in grado di rispondere alle richieste conoscitive per le quali è stato implementato, ovvero fornire dati dettagliati e tempestivi per sorvegliare le dinamiche di diffusione delle malattie nella popolazione veneziana: gli organismi preposti alla programmazione hanno a disposizione questo supporto essenziale e non possono esimersi dall'avvalersi delle informazioni di qualità da esso fornite per poter affrontare adeguatamente le esigenze degli assistiti.

Nel futuro potranno essere ulteriormente ampliati gli studi sulla popolazione veneziana condotti mediante il SEI, affinché i servizi sanitari offerti possano realmente essere efficaci ed efficienti, avendo come solido punto di partenza un quadro aggiornato dello stato di salute degli assistiti.

In generale, nell'ambito degli archivi elettronici le iniziative singole ed isolate, fondamentali per l'esplorazione di avanzamenti delle possibilità di

ricerca e per lo sviluppo di nuove metodologie, dovranno a breve trovare una sintesi a livello centrale ed istituzionale; questo non tanto per sminuire l'importanza di studi progettati e condotti autonomamente, quanto perché soltanto un loro coordinamento potrà esaltarne al meglio la forza innovativa, definendo parametri di qualità condivisi, costituendo uno spazio di confronto rigoroso di metodi e risultati, in sintesi accreditando e diffondendo strumenti all'avanguardia al servizio della salute.

BIBLIOGRAFIA

- Agenzia Italiana del Farmaco, *L'uso dei farmaci in Italia*. 2006

- Enciclopedia della medicina, DeAgostini 1995

- Ferrario M. et al., *Surveillance of ischaemic heart disease: results from the Italian MONICA populations*. International Journal of Epidemiology, Vol. 30/Suppl. 1: S23-S29, October 2001

- Gallo U. et al, *Stima della prevalenza della popolazione portatrice di patologie croniche mediante il metodo "cattura e ricattura": l'esempio del diabete mellito*. Giornale italiano di farmacia clinica, 19, 4, 2005

- Giampaoli S., Vanuzzo D., Ferrario M., Vancheri F., Cesana G., (Gruppo di Coordinamento), *Registro Nazionale degli Eventi coronarici e cerebrovascolari*, Italian Heart Journal, Vol. 5/Suppl. 3, Aprile 2004

- ISTAT, *Classificazione delle malattie, traumatismi e cause di morte*, IX revisione 1975 - Collana Metodi e Norme; serie C, n. 10

- ISTAT, *Indagine multiscopo sulle famiglie "Condizioni di salute e ricorso ai servizi sanitari" – Anno 1999 - 2000* - Collana Informazioni, n.26, 2002

- ISTAT, *Indagine multiscopo sulle famiglie "Aspetti della vita quotidiana" - Anno 2003* - Collana Informazioni, n. 25, 2005

- Istituto Superiore di Sanità, *L'integrazione di archivi elettronici per l'epidemiologia e la sanità pubblica: finalità e metodi*, Appelgren E., Luzi P. (a cura di), 2007. Collana Istisan Congressi 07/C2

- Lafisca S. et al. (a cura di), *I° Rapporto sullo stato di salute della popolazione veneziana: 1998 – 2003*

- Lauria G. et al., *Incidence and Prognosis of Stroke in the Belluno Province, Italy. First-Year Results of a Community-Based Study*, Stroke. 1995; 26: 1787-1793

- Recaldin L., *Indagine sulla salute nel Comune di Venezia attraverso un sistema di sorveglianza epidemiologica*, Relatore: Simonato L., Università degli Studi di Padova, Tesi di Laurea triennale in Statistica, Popolazione e Società, a.a. 2005-2006

- Regione del Veneto, *Indagine su inquinamento atmosferico e funzionalità respiratoria nei bambini affetti da asma bronchiale nelle AULSS 12 e 13 del Veneto*. Gennaio 2006

- R. J. Black, L. Simonato, H. H. Storm e E. Démaret (a cura di), *Automated data collection in cancer registration*, IARC Technical Reports N. 32; Lyon, 1998

- Simonato L., *Cenni di epidemiologia dei tumori*, Centro Oncologico Regionale – Registro Tumori del Veneto

- Tessari R., Canova C., Canal F. et al., *Indagine sull'inquinamento ambientale da diossine e sarcomi dei tessuti molli nella popolazione di Venezia - Mestre: un esempio di utilizzo di fonti informative elettroniche correnti*. Epidemiologia & Prevenzione, anno 30 (3) Maggio-Giugno 2006

RINGRAZIAMENTI

Ringrazio quanti mi hanno seguita ed aiutata nello svolgimento dello stage e nella stesura di questo elaborato.

In particolare, colgo l'occasione per ringraziare il Professor Lorenzo Simonato per avermi dedicato tempo ed attenzione, consigliandomi pazientemente; il Direttore del Dipartimento di Epidemiologia dell'Azienda ULSS 12 veneziana, Dottor Sergio Lafisca, per la possibilità di svolgere lo stage; le Dottoresse Cristina Canova e Roberta Tessari per il tempo dedicatomi e la disponibilità dimostratami.

Un pensiero di gratitudine alla mia famiglia, che mi ha sempre sostenuta con affetto.