

**RAPPORTO SCIENTIFICO FINALE DEL  
PROGETTO FINANZIATO DAL CENTRO  
PER IL CONTROLLO DELLE MALATTIE  
(CCM):**

**“Sviluppo di sistemi di sistemi di  
sorveglianza epidemiologica basati  
sull'utilizzo degli archivi elettronici sanitari  
in ambito di Sanità Pubblica: uno studio  
pilota in alcune ASL selezionate”**

Preparato da: Lorenzo Simonato, Cristina Canova  
Dipartimento di Medicina Ambientale e Sanità Pubblica  
Università degli Studi di Padova  
[lorenzo.simonato@unipd.it](mailto:lorenzo.simonato@unipd.it)

# Indice interattivo (cliccare sul n. di pagina)

## **Obiettivo Specifico 1** **Pag. 3**

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

## **Obiettivo Specifico 2** **Pag. 8**

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

## **Obiettivo Specifico 3** **Pag. 9**

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

**ASL 18 Rovigo** **Pag. 14**

**ASL Modena** **Pag. 47**

**ASL 2 Perugia** **Pag. 71**

## **Obiettivo Specifico 4** **Pag.119**

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

**Relazione Gruppo di Lavoro Area 1** **Pag. 123**

Strumenti e metodi per l'attivazione di database integrati presso le USL

**Relazione Gruppo di Lavoro Area 2** **Pag. 140**

Utilizzo di basi di dati integrati

**Relazione Gruppo di Lavoro Area 3** **Pag. 174**

Validazione

## **Bibliografia** **Pag.197**

## **Considerazioni finali** **Pag.211**

Vengono qui di seguito presentati in maniera dettagliata i risultati ottenuti suddivisi in rapporto agli obiettivi prefissati dal progetto:

## **OBIETTIVO SPECIFICO 1**

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

I primi 6 mesi di lavoro hanno riguardato la selezione e valutazione di alcune ASL allo scopo di organizzare presso di esse un sistema integrato delle fonti informative correnti che permettesse di individuare gli assistiti affetti da alcune patologie cronico degenerative nella popolazione generale.

In collaborazione col CNESPS sono state selezionate un numero di ASL tenendo in considerazione precedenti esperienze di collaborazione nel campo epidemiologico e delle Sanità Pubblica e cercando di rispettare una omogenea distribuzione geografica a livello nazionale. Sono state escluse a priori le ASL già coinvolte nella produzione delle stime di frequenza nel corso dei lavori del gruppo nazionale AIE-SISMEC. Le ASL non rappresentano quindi un campione né delle ASL né della popolazione italiana.

Un questionario (**allegato A**) è stato preparato ed inviato il 28/05/2008 a 9 ASL con l'obiettivo di arrivare ad includere 2-3 di queste ASL nel progetto di replicazione del Sistema Epidemiologico Integrato (SEI) già attivo a Venezia. Il programma prevedeva di verificare la disponibilità delle ASL e di verificare con una visita in loco le condizioni necessarie (esistenza ed accessibilità dei dati, personale qualificato esistente o reperibile, etc.).

Le ASL che hanno risposto al questionario sono state: Rovigo, Modena, Perugia, Barletta-Trani, Siracusa. Si è quindi proceduto alle visite in loco in occasione delle quali sono stati organizzati incontri ai quali ha partecipato il personale dei Dipartimenti Prevenzioni, le Direzioni Sanitarie, e dei Sistemi Informativi delle ASL.

Riassumiamo qui di seguito le informazioni raccolte e le conseguenti valutazioni.

**1 ASL 18 Rovigo (Veneto) pop. 235.000** (Incontro avvenuto presso il Dipartimento di Medicina Ambientale dell'Università di Padova, 11 Luglio 2008)

La situazione, per quanto riguarda gli archivi è la seguente:

- ❖ L'anagrafe sanitaria è disponibile dal 1999 ed è di tipo storico.
- ❖ Archivio dei certificati di morte dal 1996-2007
- ❖ SDO comprensive di mobilità passiva dal 1995-2007
- ❖ Prescrizioni farmaceutiche dal 1997-2007
- ❖ Anatomia Patologica dal 1995-2007

Il SEI verrebbe attivato presso il Dipartimento di Prevenzione dove c'è uno statistico part-time che potrebbe lavorare al progetto.

**2 ASL Modena (Emilia-Romagna) pop. 670.000** (visita a Modena 23 Luglio 2008)

La situazione, per quanto riguarda gli archivi è, in questa realtà, la seguente:

- ❖ L'anagrafe sanitaria è disponibile dal 2001 ed è di tipo storico in cui le informazioni sono sovrainposte temporalmente. E' anche possibile utilizzare l'anagrafe comunale limitatamente al comune di Modena (180,000 abitanti) disponibile dal 1994.
- ❖ Archivio dei certificati di morte dal 1987-2007
- ❖ SDO comprensive di mobilità passiva dal 2002 al 2007
- ❖ prescrizioni farmaceutiche dal 2002 al 2007
- ❖ esenzioni ticket dal 2002 al 2007
- ❖ Non è attualmente disponibile l'Anatomia Patologica, anche se è già attivo un Registro Tumori con dati dal 1988 al 2005
- ❖ dati di laboratorio dall'anno 2005 al 2008.

Il SEI verrebbe attivato presso il Dipartimento di Prevenzione (Servizio di Epidemiologia) diretto da Carlo Goldoni dove lavorano 4 statistici e 3 medici.

**3 ASL 8 Siracusa pop. 399.000** (visita a Siracusa 24 Luglio 2008)

La situazione, per quanto riguarda gli archivi è, in questa realtà, la seguente:

- ❖ L'anagrafe sanitaria è disponibile dal 1980 con movimenti anagrafici aggiornati mensilmente dai comuni. Una copia dell'anagrafe nei diversi periodi temporali è disponibile.

- ❖ Archivio dei certificati di morte dal 2000-2007
- ❖ SDO comprensive di mobilità passiva dal 2000 al 2006
- ❖ prescrizioni farmaceutiche dal 2004
- ❖ esenzioni ticket dal 2007
- ❖ L'Anatomia Patologica ha due fonti delle quali solo una è informatizzata dal 2000, mentre l'altra è in corso di informatizzazione.

Il SEI verrebbe attivato presso il Servizio di Epidemiologia e Medicina Preventiva diretto da Lia Contrino dove lavorano al momento solo 3 medici. E' in fase di verifica la possibilità di coinvolgere nelle attività dell'unità un informatico.

#### **4 ASL 2 Perugia (Umbria) pop. 357.000 (visita a Perugia 5 Settembre 2008)**

La situazione, per quanto riguarda gli archivi è, in questa realtà, la seguente:

- ❖ Anagrafe sanitaria disponibile dal 2005, mentre non viene attualmente utilizzata l'anagrafe comunale
- ❖ Archivio dei certificati di morte dal 1994-2006
- ❖ SDO comprensive di mobilità passiva dal 2000 al 2007
- ❖ prescrizioni farmaceutiche dal 2003 al 2007
- ❖ esenzioni ticket dal 2005 al 2007
- ❖ Non è attualmente disponibile l'Anatomia Patologica, anche se è già attivo il Registro Tumori dell'Umbria con dati dal 1994 al 2005.

Il SEI verrebbe attivato presso l'Unità Operativa Semplice di Epidemiologia diretta da Marco Petrella dove lavorano al momento 2 medici in collaborazione con il Servizio di Programmazione Controllo e Sistemi Informativi coordinato da Manuela Mollichella. E' in fase di verifica la possibilità di coinvolgere nella prima fase di linkage e di creazione di una datawarehouse la società CSIO.

#### **5 ASL BAT Barletta-Trani pop. 380.000 (visita a Barletta 15 Settembre 2008)**

Questa ASL è nata nel 2005 dalla fusione di 3 ASL. La situazione, per quanto riguarda gli archivi è, in questa realtà, la seguente:

- ❖ Anagrafe sanitaria disponibile dal 2006
- ❖ Archivio dei certificati di morte dal 2006
- ❖ SDO comprensive di mobilità passiva dal 2006
- ❖ prescrizioni farmaceutiche dal 2006
- ❖ esenzioni ticket dal 2006

❖ Non è attualmente disponibile l'Anatomia Patologica.

Il SEI verrebbe attivato presso l'Unità di Epidemiologia in staff alla Direzione Generale diretta dal Dr. Rosita priva attualmente di personale informatico-statistico in collaborazione con il Dr. Emilio Gianicolo del CNR del Salento.

In considerazione del carattere di progetto pilota e delle limitate risorse sia finanziarie che di personale disponibili viene deciso di limitare l'inclusione nel progetto delle ASL di Rovigo, Modena e Perugia pur consapevoli che tutte le ASL visitate presentano caratteristiche che potenzialmente non escludono l'attivazione del SEI in una fase successiva.

## Allegato A- questionario disponibilità archivi.

ASL di appartenenza.....

Dipartimento/Unità eventualmente responsabile del Progetto

.....

Personale statistico/informatico disponibile per il Progetto (n°/qualifica)

.....

Dimensione della popolazione assistita.....

### Disponibilità temporale archivi elettronici informatizzati completi

<b>FONTE</b>	<b>DA (ANNO)</b>	<b>A (ANNO)</b>
ANAGRAFE SANITARIA		
ANAGRAFE COMUNALE		
MORTALITA'		
SCHEDE DIMISSIONI OSPEDALIERE (SDO)		
MOBILITA' PASSIVA DELLE SDO		
PRESCRIZIONI FARMACEUTICHE		
ESENZIONI TICKET		
ANATOMIA PATOLOGICA		

## **OBIETTIVO SPECIFICO 2**

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

Sono state incontrate notevoli difficoltà nel reperire personale temporaneo ma qualificato (laurea in statistica o informatica) da impiegare nella costruzione della base dati integrata. L'offerta di borse di studio o altre forme di contratto a tempo determinato non è apparentemente competitiva e abbiamo dovuto sperimentare ben due rinunce su tre da parte di queste figure professionali. Ci siamo quindi orientati verso forme di collaborazione con ditte private con esperienza nel campo della gestione informatica di dati sanitari. Tale collaborazione è stata attuata con risultati soddisfacenti nelle ASL di Modena e di Perugia attraverso numerosi incontri dedicati a spiegare le finalità del progetto, la struttura dei dati necessari per la costruzione degli archivi. La fase di applicazione degli algoritmi e dell'analisi statistica è stata invece portata a termine a Modena da personale strutturato dell'Unità di Epidemiologia dell'ASL e a Perugia dalla ditta privata sempre usando strumenti standard forniti dal nostro Dipartimento. Nell'ASL di Rovigo è stato invece possibile arruolare tramite borsa di studio una laureata in statistica che è stata adeguatamente istruita e formata sia sotto il profilo metodologico che per quanto la riguarda l'utilizzo di softwares dedicati per effettuare le operazioni di record-linkage, l'utilizzo degli algoritmi e le analisi statistiche. I softwares utilizzati sono stati prevalentemente ACCESS e SAS per quanto riguarda gli strumenti forniti dal Dipartimento mentre le ditte private hanno utilizzato principalmente MySQL, SQL server e Stata.

## **OBIETTIVO SPECIFICO 3**

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

Questa sezione include la parte più estesa ed importante dei risultati che sono di seguito presentati suddivisi per ASL. E' stato infatti possibile in tutte tre le ASL costituire l'archivio derivante dall'incrocio delle fonti informative sanitarie elettroniche ed applicare ad esso gli algoritmi patologia – specifici.

**Periodo di svolgimento del progetto: novembre 2008 – dicembre 2009**

La realizzazione del progetto si è composta di quattro fasi:

1. sistemazione archivi
  - acquisizione degli archivi aziendali richiesti per il progetto
  - organizzazione dei dati in un unico dataset per archivio
2. valutazione della qualità dei dati
  - distribuzioni di qualità per l'Anagrafe Sanitaria
  - distribuzioni di qualità per gli altri archivi
3. record linkage degli archivi con l'Anagrafe Sanitaria
  - valutazione storicizzazione dell'archivio anagrafico
  - individuazione e valutazione delle chiavi di linkage utilizzabili per singolo archivio
  - esecuzione di procedure di record linkage semi-deterministico a cascata
  - valutazione delle percentuali di linkage
4. produzione di stime di frequenza
  - applicazione agli archivi di algoritmi forniti
  - confronto con stime ottenute applicando gli stessi algoritmi in altre realtà italiane

Gli algoritmi per la realizzazione di stime di incidenza e prevalenza di alcune patologie, applicati ai dati, corrispondono a quelli presentati nel Supplemento della Rivista "Epidemiologia e Prevenzione", Anno 32 (3), maggio-giugno 2008 e di seguito riportati schematicamente.

### Definizione casi

- Incidenza

E' stata stimata l'incidenza per l'infarto acuto del miocardio (IMA). L'algoritmo utilizzato per la selezione della casistica è il seguente:

### **Fonti dei dati e criteri di definizione di caso incidente di infarto miocardico acuto (IMA):**

Fonte	Criteri selezione casistica	Criteri definizione incidenza
mortalità	deceduto per infarto miocardico acuto in diagnosi di morte principale (ICD9-CM <sup>1</sup> : 410*; ICD10-CM <sup>1</sup> : I21*-I22*)	assenza di altri ricoveri con diagnosi di dimissione principale o secondaria con i codici ICD9-CM 410* o 412* (infarto miocardico pregresso), nei 5 anni precedenti alla data di ammissione o di morte
schede di dimissione ospedaliera	ricovero ordinario per infarto miocardico acuto (ICD9-CM <sup>1</sup> : 410*) in diagnosi di dimissione principale o secondaria, se associata ad alcuni specifici codici in diagnosi principale <sup>2</sup>	

<sup>1</sup>Classificazione internazionale delle malattie.

<sup>2</sup>Codici in diagnosi di dimissione principale: 427.1, 427.41, 427.42, 427.5, 428.1, 429.5, 429.6, 429.71, 429.79, 429.81, 518.4, 780.2, 785.51, 414.10, 423.0. Codes in principal diagnosis: 427.1, 427.41, 427.42, 427.5, 428.1, 429.5, 429.6, 429.71, 429.79, 429.81, 518.4, 780.2, 785.51, 414.10, 423.0.

- Prevalenza

E' stata stimata la prevalenza per le seguenti patologie:

- Diabete
- Cardiopatia ischemica
- BPCO
- Asma
- Malattie polmonari cronico-ostruttive (MPCO)

Sono stati applicati ai dati i seguenti algoritmi:

### **Fonte dei dati e criteri di definizione di caso prevalente di DIABETE:**

Fonte	Criteri selezione casistica	Criteri definizione prevalenza
schede di dimissione ospedaliera	dimesso nell'anno di stima o nei 4 anni precedenti con diagnosi di diabete (ICD9-CM <sup>1</sup> = 250*) in qualunque campo diagnosi, e non deceduto al 30 giugno dell'anno di stima	presenza in almeno una delle fonti
prescrizioni farmaceutiche	almeno due prescrizioni nell'anno di stima per farmaci antidiabetici (ATC <sup>2</sup> = A10*)	
esenzioni ticket	esenzione per diabete (codice esenzione = 013.250) nell'anno di stima o nei tre anni precedenti e non deceduto al 30 giugno dell'anno di stima	

<sup>1</sup>Classificazione internazionale delle malattie.

<sup>2</sup>Classificazione anatomica terapeutica chimica.

## Fonte dei dati e criteri di definizione di caso prevalente di CARDIOPATIA ISCHEMICA:

Fonte	Criteri selezione casistica	Criteri definizione prevalenza
mortalità	deceduto nell'anno di stima con diagnosi principale di morte per cardiopatia ischemica (ICD9-CM <sup>1</sup> = 410*-414*; ICD10-CM <sup>1</sup> = I20*-I25*)	presenza in almeno una delle fonti
schede di dimissione ospedaliera	dimesso nell'anno di stima o nei 4 anni precedenti con diagnosi di cardiopatia ischemica (ICD9-CM <sup>1</sup> = 410*-414*) in qualunque campo diagnosi, e non deceduto al 30 giugno dell'anno di stima	
prescrizioni farmaceutiche	almeno due prescrizioni nell'anno di stima per nitrati organici (ATC <sup>2</sup> = C01DA*)	
esenzioni ticket	esente nell'anno di stima per cardiopatia ischemica (codice esenzione = 002.414 ) nell'anno di stima o nei tre anni precedenti e non deceduto al 30 giugno dell'anno di stima	

<sup>1</sup> Classificazione internazionale delle malattie.

<sup>2</sup> Classificazione anatomica terapeutica chimica.

## Fonte dei dati e criteri di definizione di caso prevalente di BPCO:

Fonte	Criteri selezione	Criteri definizione caso prevalente
mortalità	deceduto nell'anno di stima per BPCO (codici ICD9-CM <sup>1</sup> = 490* - 492*, 494*, 496*; ICD10-CM <sup>1</sup> = J40*, J41*, J42*, J43*, J44*, J47*)	presenza in almeno una delle fonti
schede di dimissione ospedaliera	dimesso nell'anno di stima o nei quattro anni precedenti se vivi al 30 giugno dell'anno in stima, con diagnosi principale o secondaria di BPCO, (codici ICD 9-CM <sup>1</sup> = 490* - 492*, 494*, 496*)	

<sup>1</sup> Classificazione internazionale delle malattie

Codici ICD9-CM e relative voci nosologiche	
490.X	bronchite, non specificata se acuta o cronica
491.X	bronchite cronica
491	bronchite cronica semplice
491.1	bronchite cronica mucopurulenta
491.2	bronchite cronica ostruttiva
491.8	altre bronchiti croniche
491.9	bronchite cronica non specificata
492.X	enfisema
492.0	bolla enfisematosa
492.8	altri enfisemi
494.X	bronchiectasie
496.X	ostruzioni croniche delle vie respiratorie, non classificate altrove

## Fonte dei dati e criteri di classificazione di caso prevalente di ASMA:

Fonte	Criteri selezione casistica	Criteri definizione prevalenza
mortalità	Deceduto nell'anno di stima per asma (ICD9-CM= 493*; ICD10-CM= J45*, J46*) come causa di morte principale	presenza in almeno una delle fonti
schede di dimissione ospedaliera	dimesso nell'anno di stima con diagnosi di asma (ICD9-CM <sup>1</sup> = 493*) in qualunque campo diagnosi	
prescrizioni farmaceutiche	almeno una prescrizione nell'anno di stima per: <ul style="list-style-type: none"> <li>▪ Adrenergici inalatori (ATC<sup>2</sup>=R03A*)</li> <li>▪ Adrenergici per uso sistemico (ATC=R03CC02, R03CC04, R03CK)</li> <li>▪ Antagonisti dei recettori leucotrienici (ATC=R03DC01 R03DC03)</li> </ul>	
esenzioni ticket	esente nell'anno di stima per asma (codice esenzione= per 007.493)	

<sup>1</sup>Classificazione internazionale delle malattie.

<sup>2</sup>Classificazione anatomica terapeutica chimica.

## Fonte dei dati e criteri di definizione di caso prevalente di malattie polmonari cronico-ostruttive (MPCO):

Fonte	Criteri selezione casistica	Criteri definizione prevalenza
mortalità	decaduto nell'anno di stima per codici ICD9-CM <sup>1</sup> = 490*, 491*, 493*, 494* 496*; ICD10-CM=J40*, J41*, J42*, J44*, J45*, J46*, J47* come causa di morte principale	presenza in almeno una delle fonti
schede di dimissione ospedaliera	dimesso nell'anno di stima con codice ICD9-CM <sup>1</sup> = 490*-494*, 496* in qualunque campo diagnosi	
prescrizioni farmaceutiche	almeno una prescrizione nell'anno di stima per farmaci respiratori: <ul style="list-style-type: none"> <li>▪ Adrenergici inalatori (ATC<sup>2</sup>=R03A*)</li> <li>▪ Adrenergici per uso sistemico (ATC=R03CC02, R03CC04, R03CK)</li> <li>▪ Antagonisti dei recettori leucotrienici (ATC=R03DC01 R03DC03)</li> <li>▪ Anticolinergici (ATC=R03BB01, R03BB02)</li> </ul> Derivati xantini: ATC=R03DA01, R03DA04, R03DA05, R03DA08, R03DA11, R03DA49	
esenzioni ticket	esente nell'anno di stima per asma (codice=007) e/o insufficienza respiratoria (codice=024)	

<sup>1</sup>Classificazione internazionale delle malattie.

<sup>2</sup>Classificazione anatomica terapeutica chimica.

## Popolazioni

Nel calcolo dei tassi annuali, al denominatore si è utilizzata la popolazione al 30 giugno dell'anno, ricavata da fonte ISTAT ([www.demo.istat.it](http://www.demo.istat.it)).

## Standardizzazione

Il metodo di standardizzazione è diretto per classi d'età quinquennali e utilizza un'unica popolazione.

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 e più":

<b>CLASSI DI ETÀ</b>	<b>Totale</b>
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
<b>Totale</b>	<b>56 995</b>

### Contributi

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  $\geq 100\%$ .

Di seguito verrà illustrata in dettaglio ogni fase per le 3 ASL selezionate nel progetto, separatamente.

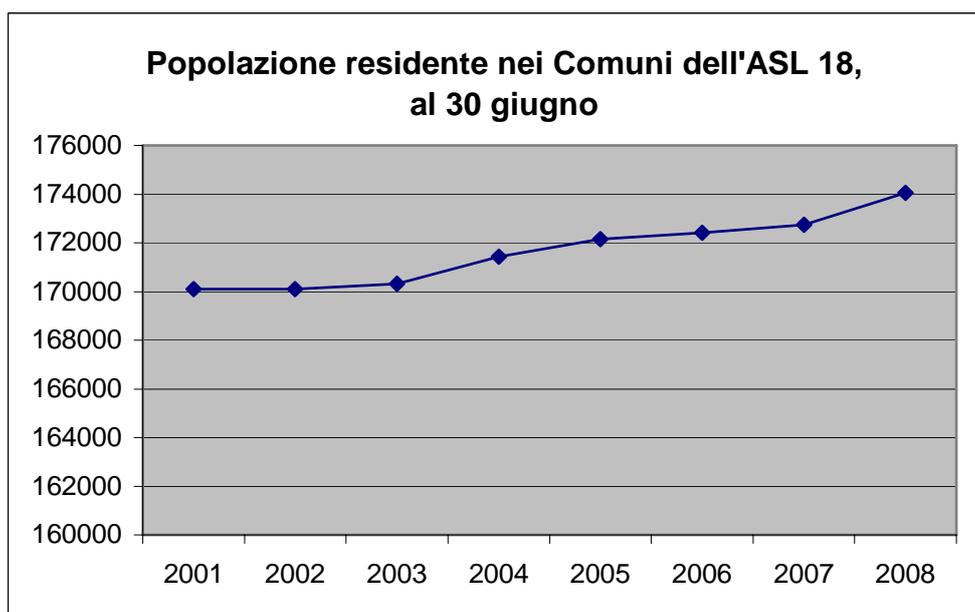
## ASL 18 Rovigo

Il territorio dell'Azienda Sanitaria Locale 18 di Rovigo comprende 41 Comuni, tra i quali Rovigo, capoluogo di provincia.

La seguente tabella riporta la popolazione residente in questi 41 Comuni al 30 giugno degli anni dal 2001 al 2008, per i quali è stato possibile produrre stime di frequenza:

anno	totale
2001	170106
2002	170106
2003	170323
2004	171428
2005	172151
2006	172420
2007	172737
2008	174045

Fonte: ISTAT ([www.demo.istat.it](http://www.demo.istat.it))



Si evidenzia una tendenza all'aumento della popolazione, più marcata nell'anno 2008; nell'intero periodo l'aumento si attesta a 4000 persone.

## **Fase 1: sistemazione archivi**

Tempo di realizzazione: 1 novembre - 15 dicembre 2008

### **1.1 - Acquisizione degli archivi aziendali richiesti per il progetto**

Gli archivi dell'ASL 18 di Rovigo disponibili al 1 gennaio 2009 sono i seguenti:

**Tab. 1.1:** Archivi ASL 18 richiesti per il progetto; sono specificati primo e ultimo anno disponibile ed il formato dei dati

<b>FONTE</b>	<b>DA (anno)</b>	<b>A (anno)</b>	<b>FORMATO</b>
ANAGRAFE SANITARIA	1996	2008	testo, lunghezza fissa
ANAGRAFE COMUNALE	archivio non disponibile		
MORTALITA'	1987	2008	excel
SCHEDE DIMISSIONI OSPEDALIERE (SDO)	1995	2008	mdb
MOBILITA' PASSIVA DELLE SDO	2000	2008	mdb
PRESCRIZIONI FARMACEUTICHE	1997	2008	mdb
ESENZIONI TICKET	1991	2008	testo, lunghezza fissa
ANATOMIA PATOLOGICA	1995	2008	testo, lunghezza fissa

Tutti gli archivi sono stati acquisiti ed organizzati in un database utilizzando Microsoft Access.

### **1.2 - Organizzazione dei dati in un unico dataset per archivio**

Gli archivi, ad eccezione delle Esenzioni Ticket, sono stati ottenuti in file separati per anno.

Si è riscontrata in generale una significativa variabilità nei nomi delle variabili e nella quantità delle stesse per anno; questo ha richiesto un lavoro di uniformazione prima della costituzione di un unico database e, più volte, è stato necessario chiedere spiegazioni direttamente ai gestori dei dati per comprendere alcune difformità.

La sistemazione di ogni archivio ha richiesto un lavoro differente, in quanto sono emerse problematiche diverse.

Alcune problematiche specifiche riscontrate a Rovigo:

- Date di nascita inferiori al 1 gennaio 1900 non erano accettate da Access in fase di importazione da Excel nel formato gg/mm/aaaa;
- SDO 1997-98: Nome e Cognome si trovano nello stesso campo, separati da \*; bisogna tenere conto anche di questa possibilità nella definizione di procedure il più possibile automatizzate
- I nomi delle variabili cambiano spesso di anno in anno, talvolta con l'aggiunta di alcune nuove variabili: è necessario capire quali variabili corrispondano, anche facendo riferimento alla normativa regionale
- I cicli di prestazioni day hospital dovrebbero essere conteggiati come unica SDO, annotando il numero complessivo di accessi. E' stato riscontrato per alcuni anni che venga specificato ogni singolo accesso; si trovano quindi molti record ripetuti, uno per ogni accesso, e si rende necessario ricostruire l'informazione standardizzando con gli altri anni disponibili.

## **Fase 2: valutazione della qualità dei dati**

Tempo di realizzazione: 1 gennaio - 15 febbraio 2009

### **2.1 - Distribuzioni di qualità per l'Anagrafe Sanitaria**

La chiave dell'archivio è la variabile "Tessera Sanitaria".

Per quanto riguarda campi importanti per il record linkage e per analisi successive, è stata valutata la completezza annuale come percentuale dei campi compilati sul totale dei record.

La completezza del campo "Codice Fiscale" è attorno all'80%, inferiore a quanto atteso; si è ritenuto opportuno applicare un algoritmo di calcolo del CF sulla base dei dati anagrafici disponibili per ogni soggetto, aggiungendo al dataset il CF ricalcolato come nuova variabile.

Di seguito si riporta la corrispondenza fra CF e CF ricalcolato per l'anno 2008:

**Tab. 2.1:** Corrispondenza fra CF originale e CF ricalcolato – Anagrafe Sanitaria 2008

	<b>SI</b>	<b>NO</b>	<b>CF originario nullo</b>	<b>nr tot record</b>
Corrispondenza CF	196023	10660	50022	256705
% sul totale dei record	76,4	4,2	19,5	100,0

Il CF ricalcolato corrisponde al CF originale nel 94,8% dei record.

E' stata calcolata la completezza anche dei campi "Cognome", "Nome", "Data di nascita", "Indirizzo" di residenza, che è del 100%.

Nel corso dei controlli sono emersi assistiti duplicati, ovvero soggetti con diversa Tessera Sanitaria, ma con stesso CF o più variabili anagrafiche.

Si è approfondita la ricerca dei duplicati e sono stati considerati tali i record con corrispondenza completa in una delle seguenti combinazioni di variabili anagrafiche:

1. codice fiscale, cognome, nome, data di nascita
2. codice fiscale, cognome, data di nascita
3. codice fiscale, cognome, nome
4. codice fiscale, nome, data di nascita
5. nome, cognome, data di nascita

I soggetti duplicati (153 record nell'Anagrafe 2007) sono stati analizzati manualmente per cercare di identificare i duplicati reali ed eventuali altri problemi associati a questi record.

Si è quindi deciso di mantenere, fra due duplicati, il record con il progressivo di TS più alto (più recente). E' mantenuta una copia dell'Anagrafe integrale; nel nuovo database è stata aggiunta una variabile per tracciare le modifiche effettuate.

L'elenco dei soggetti duplicati è stato inviato all'ufficio di gestione dell'Anagrafe Sanitaria per

procedere a controlli interni.

Si è ritenuto di valutare l'attendibilità del campo "Posizione assistito", per il quale si riportano legenda e distribuzione per gli ultimi 4 anni disponibili:

**Tab. 2.2:** Legenda codici di sospensione

<b>Codice</b>	<b>Descrizione</b>
Ø	ASSISTITO ATTIVO con o senza scelta del MMG/PLS
1	TRASFERIMENTO
2	SCADENZA ISCRIZIONE
3	DECESSO
4	SERVIZIO MILITARE
5	RECUSAZIONI ECCESSO MASSIMALE
6	ANNULLO
7	MEDICO CHE RICUSA
8	ISCRIZIONE PROVVISORIA IN ALTRA USL
9	CASI DI LEGGE
10	DOPPIA ISCRIZIONE
11	X VARIAZ.DATI ANAG.E ASSEGNAZ.NUOVA TESS
14	SOGGIORNO TEMPORANEO ALL'ESTERO
22	PERDITA DIRITTO
23	ACCORPAMENTO COMUNI NUOVI DISTRETTI
24	CESSAZIONE T.P. PER TESSERA DEFINITIVA
25	CESSAZIONE RAPPORTO DI FIDUCIA
26	USCITA RESIDENZA PROTETTA
27	CAMBIO DISTRETTO DI RESIDENZA
28	VARIAZ.MEDICO IN RESID.PROTETTA
29	REVOCA RINNOVO PERMESSO DI SOGGIORNO
30	IRREPERIBILE ALL'ANAGRAFE COMUNALE
50	RIASSEGNAZIONE CODICE TESSERA
80	STESSO COD.SAN. ASS.TO ALTRA ULSS VENETO
88	VOLONTA' DELL'ASSISTITO
89	AUT. SCELTA FUORI PROVINCIA
90	ALTRE
95	CESSAZIONE DEL MEDICO
96	DEROGA OUT
97	NUOVA SCELTA
99	CESSATA ETA' PEDIATRICA

**Tab. 2.3:** Distribuzione record per anno e codice di sospensione dell'assistenza – Anagrafe Sanitaria

COD	2005		2006		2007		2008	
	freq	%	freq	%	freq	%	freq	%
<b>0</b>	170519	70,0	166175	67,5	162272	64,6	158942	61,9
<b>1</b>	31650	13,0	31416	12,8	32333	12,9	33641	13,1
<b>2</b>	4595	1,9	5010	2,0	6013	2,4	7242	2,8
<b>3</b>	34377	14,1	37194	15,1	39722	15,8	42106	16,4
<b>4</b>	12	0,0	7	0,0	6	0,0	4	0,0
<b>5</b>	9	0,0	278	0,1	461	0,2	578	0,2
<b>6</b>	39	0,0	61	0,0	85	0,0	625	0,2

7	147	0,1	143	0,1	141	0,1	104	0,0
8	537	0,2	769	0,3	969	0,4	1206	0,5
9	37	0,0	2901	1,2	5064	2,0	6572	2,6
10	0	0,0	10	0,0	242	0,1	260	0,1
11	3	0,0	2	0,0	1	0,0	1	0,0
14	166	0,1	143	0,1	140	0,1	131	0,1
15	0	0,0	1	0,0	1	0,0	1	0,0
22	4	0,0	2	0,0	2	0,0	2	0,0
23	0	0,0	0	0,0	0	0,0	0	0,0
24	0	0,0	65	0,0	114	0,0	159	0,1
25	0	0,0	0	0,0	0	0,0	0	0,0
26	15	0,0	4	0,0	4	0,0	3	0,0
27	2	0,0	0	0,0	0	0,0	0	0,0
28	0	0,0	0	0,0	0	0,0	0	0,0
29	0	0,0	0	0,0	0	0,0	15	0,0
30	584	0,2	714	0,3	849	0,3	1106	0,4
50	0	0,0	0	0,0	0	0,0	0	0,0
80	0	0,0	51	0,0	160	0,1	174	0,1
82	0	0,0	0	0,0	974	0,4	1893	0,7
84	0	0,0	334	0,1	716	0,3	1140	0,4
85	0	0,0	94	0,0	2	0,0	484	0,2
88	50	0,0	16	0,0	10	0,0	9	0,0
89	21	0,0	18	0,0	18	0,0	13	0,0
90	590	0,2	556	0,2	541	0,2	4	0,0
95	249	0,1	223	0,1	278	0,1	205	0,1
96	1	0,0	1	0,0	1	0,0	1	0,0
97	0	0,0	1	0,0	1	0,0	1	0,0
99	85	0,0	97	0,0	88	0,0	83	0,0
<b>tot</b>	<b>243692</b>		<b>246286</b>		<b>251208</b>		<b>256705</b>	

nota: negli anni 2006, 2007 e 2008 vi sono alcuni codici aggiuntivi (15, 82, 84, 85) rispetto alla legenda

## 2.2 - Distribuzioni di qualità per gli altri archivi

Per valutare eventuali evidenti discordanze numeriche è stato annotato il totale dei record per anno per i singoli archivi.

**Tab. 2.4:** Distribuzione record per archivio, per anno di disponibilità

ANNO	Mortalità	SDO	Mobilità Passiva	Farmaceutica	Esenzioni Ticket	Anatomia Patologica
1987	2406					
1988	2245					
1989	1043					
1990	2231					
1991	2213				18138	
1992	2271				6791	
1993	2217				17595	
1994	2356				7747	
1995	2072	62931			5784	51901
1996	2121	79318			4601	50700
1997	2168	52149		1031019	4880	57062
1998	2187	48534		1064111	4997	64301
1999	2276	42488		1138119	13707	84708
2000	2257	38433	8925	1197358	25159	92304
2001	2192	35648	7458	1407758	4293	84993
2002	2222	34989	9019	1482100	8632	88114
2003	2230	32836	9316	1388218	20524	89020
2004	2192	31153	9768	1458860	17716	105690
2005	2263	33893	9753	1470170	10232	100089
2006	2245	33465	9914	1529184	8223	116062
2007	2173	32075	9827	1609873	8420	98044(*)
2008	2137	30741	9296	1666282	2623	

(\*) incompleto

Per i singoli archivi sono state fatte anche valutazioni di qualità più specifiche.

### Mortalità

Poiché per l'applicazione degli algoritmi di stima sono necessarie più fonti relative ad uno stesso periodo, l'archivio di mortalità è stato considerato a partire dall'anno 1996; a questi dati di mortalità si farà riferimento nella trattazione che segue.

**Tab. 2.5:** Distribuzione record per anno e mese - mortalità

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
gen	195	262	196	257	229	204	223	214	232	220	215	223	200
feb	195	157	180	255	213	183	192	215	195	255	173	196	242
mar	217	188	224	207	194	193	207	178	205	223	225	215	187
apr	205	186	160	175	175	175	194	178	188	191	201	176	157
mag	165	189	185	190	162	201	175	179	192	171	161	168	159
giu	170	155	157	151	208	174	208	193	139	150	185	179	161
luglio	175	158	183	176	158	171	183	190	154	164	188	151	158
ago	149	135	162	155	174	158	175	193	193	158	155	168	172
set	167	165	164	175	164	182	142	152	169	154	169	140	174
ott	157	173	182	169	164	159	173	185	171	171	198	190	183
nov	180	201	186	163	200	174	166	161	170	209	171	195	190
dic	146	199	208	203	216	218	184	192	184	197	204	172	154
tot	2121	2168	2187	2276	2257	2192	2222	2230	2192	2263	2245	2173	2137

Si osserva una distribuzione nel complesso uniforme dei record mensili, come atteso.

Dal confronto del totale dei record per anno e dalle date di morte registrate negli archivi, erano inizialmente emerse alcune incompletezze macroscopiche: è stato necessario riottenere i dati originali, corretti.

E' stata inoltre valutata la distribuzione dei record di mortalità per grandi gruppi di cause di morte, per anno (vedasi **Allegato B.1**).

**Tab. 2.6:** Distribuzione record per gruppo di causa di morte e anno - mortalità

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	tot
<b>1996</b>	4	619	54	7	29	32	961	80	92	16	1	0	13	7	3	73	130	2121
<b>1997</b>	3	610	52	14	37	39	969	101	119	22	0	1	17	3	5	65	111	2168
<b>1998</b>	10	710	64	7	23	34	971	107	88	24	0	0	5	0	0	40	104	2187
<b>1999</b>	15	703	82	8	65	43	926	143	102	26	0	4	3	5	1	25	125	2276
<b>2000</b>	5	731	86	8	56	58	890	126	119	28	0	4	18	6	1	14	107	2257
<b>2001</b>	11	679	58	9	57	53	917	126	98	30	0	5	16	2	0	26	105	2192
<b>2002</b>	11	702	61	12	39	50	971	124	101	23	0	0	14	5	0	28	81	2222
<b>2003</b>	7	680	70	9	50	51	961	135	105	28	0	3	12	1	0	26	92	2230
<b>2004</b>	16	704	75	7	58	63	866	159	97	23	0	4	5	5	0	18	92	2192
<b>2005</b>	27	599	67	9	66	66	1041	153	93	20	0	1	8	2	0	11	100	2263
<b>2006</b>	12	734	86	10	89	59	882	140	85	28	0	4	5	8	0	4	99	2245
<b>2007</b>	13	594	78	9	75	48	856	124	83	35	0	5	5	2	0	8	95	2030
<b>tot</b>	134	8065	833	109	644	596	11211	1518	1182	303	1	31	121	46	10	338	1241	26383

Il numero di morti per anno per le diverse cause è piuttosto uniforme, come atteso.

Nell'anno 2008 è riportata nell'archivio la sola codifica ICD – X.

#### Schede di dimissione ospedaliera (SDO)

Nell'archivio SDO i nomi delle variabili cambiano di frequente di anno in anno, spesso con l'aggiunta di alcune nuove variabili: è stato necessario capire quali variabili fossero corrispondenti, anche facendo riferimento alla normativa regionale.

Fino al 1995 l'Azienda comprendeva anche la popolazione successivamente assistita dall'ASL 19: per questo motivo il numero di record nei primi 2 anni riportati è così elevato (rispettivamente 62931 nel 1995, 79318 nel 1996).

Fino al 1999 l'archivio comprende anche la mobilità passiva.

Anche per questo archivio è stata valutata la distribuzione per anno dei record per grandi gruppi di malattie secondo la classificazione ICD - IX (vedasi **Allegato B.1**).

**Tab. 2.7:** Distribuzione record per gruppo di causa e anno – SDO

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	codici V	tot
<b>1997</b>	1466	7656	2902	518	1328	5192	7579	4230	4171	4640	2372	1259	2325	493	261	2447	3083	216	52138
<b>1998</b>	1063	7791	2019	483	1116	5091	6899	3976	3644	3943	2394	1103	2184	372	373	2156	3086	835	48528
<b>1999</b>	694	6432	665	396	1007	4180	7132	3813	3801	3353	2136	620	1757	222	455	1674	2418	1733	42488
<b>2000</b>	459	4799	518	470	781	3597	6501	3783	3475	2930	1953	498	1800	201	210	1486	2094	2875	38430
<b>2001</b>	448	4252	581	481	659	3342	6364	3045	3344	2936	1791	470	1716	219	240	1359	1922	2479	35648
<b>2002</b>	447	4246	562	489	622	3397	6592	2888	3285	2887	1884	444	1804	158	215	599	1869	2601	34989
<b>2003</b>	389	3710	544	481	658	2482	6589	2782	3148	2783	1771	394	1887	195	245	396	1802	2580	32836
<b>2004</b>	382	3747	608	452	590	1709	6425	2655	3120	2543	1871	567	1922	175	216	304	1557	2310	31153
<b>2005</b>	421	3940	618	502	450	2922	7201	2812	3469	2480	1889	555	2152	256	231	312	1617	2066	33893
<b>2006</b>	435	4141	628	493	365	2917	6969	2795	2931	2430	1851	457	2415	217	192	595	1492	2142	33465
<b>2007</b>	504	3995	515	489	393	2191	6048	2790	2905	2474	1800	384	2693	193	144	490	1536	2531	32075
<b>2008</b>	490	3909	419	424	463	1922	5534	2618	2432	2338	1981	387	2809	225	148	634	1471	2537	30741
<b>tot</b>	6708	54709	10160	5254	7969	37020	74299	35569	37293	33399	21712	6751	22655	2701	2782	11818	22476	22368	446384

## Mobilità passiva delle SDO

Negli archivi di mobilità passiva intra-regionale sono stati riscontrati numerosi record duplicati (in totale 750 dal 2000 al 2007), che differivano soltanto per il campo "Posizione contabile"; poiché l'informazione non interessa, è stato tenuto un solo record, ovvero quello più completo nei campi "Diagnosi principale", "Tessera Sanitaria" e "Codice Fiscale", nell'ordine.

**Tab. 2.8:** Distribuzione record per anno e provenienza (extra- o intra-regionale) – mobilità passiva SDO

anno	nr record		
	extra	intra	totale
2000	4050	4875	8925
2001	4136	3322	7458
2002	3954	5065	9019
2003	4148	5168	9316
2004	4146	5622	9768
2005	4281	5472	9753
2006	4468	5446	9914
2007	4625	5202	9827
2008	4575	4721	9296

La distribuzione è, nel complesso, omogenea.

## Prescrizioni farmaceutiche

**Tab. 2.9:** Distribuzione record per anno e mese – Prescrizioni farmaceutiche

	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
gen	97563	91356	98849	108185	121032	141296	128576	128574	124800	131431	146450	155510
feb	85618	86429	95259	99282	110498	123521	115009	115729	122932	123247	128363	139924
mar	88295	93639	102468	110151	122850	130965	121423	131234	128179	137752	140913	136667
apr	92573	89404	93391	94615	114296	124796	111315	122739	120948	118621	128276	140370
mag	90060	86483	97406	107535	127279	131202	119510	123282	126501	137283	140402	143827
giu	81962	86232	92856	98265	119184	113304	109310	118250	121440	127661	130703	132115
luglio	83034	85649	89887	91074	111632	122841	110610	119063	113434	118300	127841	138049
ago	74253	79402	86413	91861	109762	112382	100954	110484	116258	121790	126221	122065
set	84300	88149	91881	94440	105751	117213	114649	118077	120361	121497	124661	136706
ott	87307	92808	93628	102225	127199	133176	123099	121552	122658	129505	143250	146139
nov	79617	90373	96040	100624	123869	114432	110500	122486	125386	134201	141215	133980
dic	86437	94187	100041	99101	114406	116972	123263	127390	127273	127896	131578	140930
tot		1064111	1138119	1197358	1407758	1482100	1388218	1458860	1470170	1529184	1609873	1666282

Nel complesso non si evidenziano aspetti problematici.

E' stata valutata la distribuzione per anno dei record per classe ATC di prescrizione farmaceutica (vedasi **Allegato B.2**):

**Tab. 2.10:** Distribuzione record per classe ATC e anno – Prescrizioni farmaceutiche

	A	B	C	D	G	H	J	L	M	N	P	R	S	V	nr	tot
97	137911	36500	366734	13046	63238	19988	144718	7971	81813	45539	2916	91177	19045	423		1031019
98	142680	40171	388082	14300	62157	21745	143907	7461	78718	47049	3003	93600	20962	276		1064111
99	146981	53052	423778	18659	59647	28030	147504	7663	73745	52801	3249	96877	25848	285		1138119
00	155212	61464	450407	20677	59136	30605	139645	8442	81234	58862	3124	99774	28393	383		1197358
01	184786	77450	515742	25960	63096	44079	148833	9174	105042	81719	3254	113062	35177	384		1407758
02	197665	85588	587232	21080	62258	44998	142318	9797	102322	89763	3391	100834	34545	309		1482100
03	177822	90073	589027	6099	57146	39432	139933	10005	95120	83306	3088	72015	24752	400		1388218
04	199921	98726	604815	5979	52888	41427	140207	10456	105060	88623	3212	81542	25647	357		1458860

<b>05</b>	213135	102549	593030	6165	48771	45597	145615	11001	98794	91859	3407	84323	25609	315		1470170
<b>06</b>	237999	108445	607814	6602	47547	50096	142380	12272	103023	99424	3733	83189	26363	297		1529184
<b>07</b>	265159	119671	620083	6365	46072	52718	153688	12863	108471	105203	4052	88743	26374	409	2	1609873
<b>08</b>	292264	125872	623935	6869	46074	54362	150595	13190	120520	114026	4181	86709	27033	647	5	1666282
<b>tot</b>	2351535	999561	6370679	151801	668030	473077	1739343	120295	1153862	958174	40610	1091845	319748	4485	7	14776770

La distribuzione mostra complessivamente una tendenza all'aumento delle prescrizioni.

La principale eccezione è data dalla classe D dei farmaci dermatologici; alla luce di accertamenti più approfonditi questa diminuzione non sembra essere attribuibile a problemi di qualità dei dati.

### Esenzioni Ticket (ET)

Si osserva notevole variabilità nel numero di record per anno, che può essere dovuta a variazioni nella normativa che regola le ET. Anche i codici di esenzione per alcune patologie mostrano marcata variabilità nel periodo.

### Anatomia Patologica (AP)

Per quanto riguarda gli ultimi 2 anni, il numero di record non è in linea con gli anni precedenti. Sono stati richiesti nuovamente i dati al Registro Tumori di Padova, ma la discrepanza permane e non ne è ancora stata chiarita la motivazione.

Questo archivio, comunque, non era richiesto nel presente progetto.

## **Fase 3: record linkage degli archivi con l'Anagrafe Sanitaria**

**Tempo di realizzazione: maggio – settembre 2009**

### *3.1 - Valutazione storicizzazione dell'archivio anagrafico*

Poiché gli archivi anagrafici sono storicizzati, ci si attende che i soggetti presenti nell'AS 2007 siano presenti anche nell'AS 2008.

Sono possibili, invece, nuovi assistiti nel 2008 non presenti nel 2007 (nuovi nati, immigrati). Gli assistiti nel 2007 non presenti nell'archivio 2008 dovrebbero limitarsi alla correzione di

eventuali errori.

Per verificare queste ipotesi sono state incrociate le AS 2005, 2006 e 2007 con l'AS 2008:

**Tab. 3.1:** Incrocio AS 2005, 2006 e 2007 con AS 2008

<b>Anno anagrafe</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>
<b>totale record</b>	243692	246286	251208
<b>nr record linkati<sup>^</sup> con AS 2008</b>	239979	244747	250429
<b>nr record non linkati<sup>^</sup></b>	3713	1539	779
<b>% record non linkati<sup>^</sup></b>	1,5%	0,6%	0,3%
<b>nr record linkati<sup>^^</sup> con AS 2008</b>	241804	245917	251041
<b>nr record non linkati<sup>^^</sup></b>	1888	369	167
<b>% record non linkati<sup>^^</sup></b>	0,8%	0,1%	0,1%

<sup>^</sup> utilizzando TS, nome, cognome, data di nascita

<sup>^^</sup> utilizzando solo TS

Per approfondire il confronto fra le due Anagrafi più recenti è stata incrociata l'AS 2006 e 2007 con l'AS 2008 per i seguenti passi di chiave, gli stessi utilizzati per l'incrocio fra Anagrafe e altre fonti:

step 1: tessera sanitaria, codice fiscale, cognome, nome, data di nascita

step 2: tessera sanitaria, cognome, nome, data di nascita

step 3: tessera sanitaria, cognome, data di nascita

step 4: tessera sanitaria, cognome, nome

step 5: tessera sanitaria, nome, data di nascita

step 6: codice fiscale, cognome, nome, data di nascita

step 7: codice fiscale, cognome, data di nascita

step 8: codice fiscale, cognome, nome

step 9: codice fiscale, nome, data di nascita

step 10: cognome, nome, data di nascita

step 11: tessera sanitaria

**Tab. 3.2:** Incrocio AS 2006 e 2007 con AS 2008 – dettaglio per passo di chiave

Passo di chiave	2007				2006			
	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
K0	251208				246286			
TCFCNDN	250178	250178	99,59	99,59	241558	241558	98,08	98,08
TCNDN	251	250429	0,10	99,69	3189	244747	1,29	99,38
TCDN	406	250835	0,16	99,85	844	245591	0,34	99,72
TCN	108	250943	0,04	99,89	173	245764	0,07	99,79
TNDN	77	251020	0,03	99,93	116	245880	0,05	99,84
CFCNDN	2	251022	0,00	99,93	30	245910	0,01	99,85
CFCDN	3	251025	0,00	99,93	2	245912	0,00	99,85
CFCN	0	251025	0,00	99,93	2	245914	0,00	99,85
CFNDN	1	251026	0,00	99,93	1	245915	0,00	99,85
CNDN	2	251028	0,00	99,93	19	245934	0,01	99,86
TS	20	251048	0,01	99,94	37	245971	0,02	99,87
NIENTE	160	251208	0,06	100,00	315	246286	0,13	100,00

**Legenda:**

T: Tessera Sanitaria

CF: Codice Fiscale

C: Cognome

N: Nome

DN: Data di Nascita

 record che nell'incrocio degli archivi con l'Anagrafe 2008 verranno comunque linkati

 record che nell'incrocio degli archivi con l'Anagrafe 2008 non verranno linkati

 record eliminati da un anno all'altro

Dopo lo step 1, per l'AS 2007 risultano non incrociati 1030 record, pari allo 0,41% dei record, mentre per l'AS 2006 non sono appaiati 4728 record, pari all'1,92%.

Al termine dell'incrocio mediante i primi 10 passi di chiave risultano non incrociati 160 record

nel 2007, 315 nel 2006, rispettivamente pari allo 0,06% e allo 0,13% del totale.

La frequenza maggiore di modifiche si ha nel nome (si tratta probabilmente di correzione di errori).

Alla luce di queste considerazioni, si è deciso di agganciare gli archivi fino al 2007 all'Anagrafe 2007, gli archivi del 2008 all'Anagrafe 2008.

### 3.2 - Individuazione e valutazione delle chiavi di linkage utilizzabili per singolo archivio

Prima di procedere al record linkage fra archivi ed Anagrafe Sanitaria sono state individuate le variabili da utilizzare nel linkage, per impostare l'algoritmo specifico per ogni archivio; si è quindi valutata la completezza dei campi stessi.

Le chiavi di linkage individuate nel complesso sono le seguenti:

- ❖ **Tessera sanitaria (TS)**
- ❖ **Codice fiscale (CF)**
- ❖ **Nome**
- ❖ **Cognome**
- ❖ **Data di nascita**

Si è reso necessario uniformare i formati delle variabili utilizzate come chiavi di linkage con il formato delle medesime nell'archivio anagrafico.

Ad esempio, il campo "Tessera Sanitaria" nell'archivio anagrafico è numerico, mentre negli altri archivi è in formato testo. Il linkage richiede formati corrispondenti fra le variabili. Poiché l'Anagrafe Sanitaria è il riferimento per il linkage, negli altri dataset è stato necessario creare una nuova variabile trasformando la TS in formato numerico, mediante la creazione di un modulo apposito in Access.

Sono stati valutati per ogni fonte la completezza, per anno e nel totale, delle singole chiavi di linkage ed il loro formato.

### Mortalità

**Tab. 3.3:** Percentuale di completezza chiavi di linkage per anno - mortalità

chiavi	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>CF</b> formato testo	0,0	0,0	4,3	95,9	99,4	99,6	99,8	98,7	99,7	100	98,8	100	99,9
<b>nome</b> formato testo	100	100	100	100	100	100	100	100	100	100	100	100	100
<b>cognome</b> formato testo	100	100	100	100	100	100	100	100	100	100	100	100	100
<b>data nascita</b> formato data/ora	100	100	100	100	100	100	100	100	100	100	100	100	100

### Schede di dimissione ospedaliera (SDO)

**Tab. 3.4:** Percentuale di completezza chiavi di linkage per anno – SDO

chiavi	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato numerico, precisione doppia	76,0	72,3	73,8	72,4	75,3	76,2	76,3	73,8	69,6	85,8	76,8	75,9
<b>CF</b> formato testo	23,5	27,3	23,8	93,8	96,0	97,0	99,4	99,8	99,8	99,0	100	100
<b>nome</b> formato testo	98,2	99,3	100	100	100	100	100	100	100	100	100	100
<b>cognome</b> formato testo	98,2	99,3	100	100	100	100	100	100	100	100	100	100
<b>data nascita</b> formato data/ora	100	100	100	100	100	100	100	100	100	100	100	100

## Mobilità passiva delle SDO

**Tab. 3.5:** Percentuale di completezza chiavi di linkage per anno – mobilità passiva intra-regionale delle SDO

chiavi	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato numerico, precisione doppia	93,5	94,1	92,6	90,1	90,8	87,6	48,2	46,8	48,6
<b>CF</b> formato testo	88,4	94,3	94,4	94,6	94,5	91,1	98,3	99,5	99,6
<b>nome</b> formato testo	99,9	100	100	100	100	100	100	100	99,9
<b>cognome</b> formato testo	100	100	100	100	100	100	100	100	100
<b>data nascita</b> formato data/ora	99,9	100	100	100	100	100	100	100	100

**Tab. 3.6:** Percentuale di completezza chiavi di linkage per anno – mobilità passiva extra-regionale delle SDO

chiavi	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato numerico, precisione doppia	3,7	6,7	3,3	3,5	3,2	2,4	2,7	2,1	1,1
<b>CF</b> formato testo	96,1	95,5	97,2	97,5	98,5	96,6	96,4	99,4	99,7
<b>nome</b> formato testo	70,9	92,1	77,6	95,6	89,8	95,4	94,2	77,2	78,9
<b>cognome</b> formato testo	70,9	92,1	77,6	95,6	89,8	95,4	94,2	77,2	78,9
<b>data nascita</b> formato data/ora	99,9	99,9	100	100	100	100	100	100	100

## Prescrizioni farmaceutiche

**Tab. 3.7:** Percentuale di completezza chiavi di linkage per anno – Prescrizioni farmaceutiche

chiavi	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato numerico, precisione doppia	99,8	99,5	99,8	99,9	99,9	99,9	99,8	99,9	98,8	97,7	97,8	98,2
<b>CF</b> formato testo	0,2	0,5	0,2	0,1	0,1	0,1	0,2	0,1	1,2	2,3	2,2	1,8
<b>nome</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>cognome</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>data nascita</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0

**nota:** in questo archivio, CF e TS sono contenuti in un unico campo: ad ogni prescrizione nell'archivio corrisponde pertanto una sola variabile anagrafica relativa al soggetto assistito

## Esenzioni Ticket (ET)

**Tab. 3.8:** Percentuale di completezza chiavi di linkage per anno – Esenzioni Ticket

chiavi	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato numerico, precisione doppia	100	100	100	100	100	100	100	100	100
<b>CF</b> formato testo	98,3	99,7	99,2	99,6	99,6	99,5	99,2	98,8	n.d.
<b>nome</b> formato testo	100	100	100	100	100	100	100	100	100
<b>cognome</b> formato testo	100	100	100	100	100	100	100	100	100
<b>data nascita</b> formato data/ora	100	100	100	100	100	100	100	100	100

## Anatomia Patologica (AP)

**Tab. 3.9:** Percentuale di completezza chiavi di linkage per anno – Anatomia Patologica

chiavi	2000	2001	2002	2003	2004	2005	2006
<b>TS</b> formato numerico, precisione doppia	96,0	96,5	96,0	96,3	96,0	93,8	91,1
<b>CF</b>	0,0	0,0	0,0	0,0	0,0	0,0	0,0
<b>nome</b> formato testo	100	100	100	100	100	100	100
<b>cognome</b> formato testo	100	100	100	100	100	100	100
<b>data nascita</b> formato data/ora	100	100	100	100	100	100	100

La procedura utilizzata per il record linkage è di tipo semi-deterministico, ovvero un record è considerato riconducibile al medesimo soggetto solo se la corrispondenza fra campi riferiti alla stessa variabile è perfetta, con una procedura a cascata di passi sempre meno restrittivi.

Le informazioni anagrafiche contenute in tutti gli archivi sono state utilizzate per identificare i soggetti nell'Anagrafe mediante i seguenti 5 passi di chiave:

- step 1: tessera sanitaria, cognome, nome, data di nascita
- step 2: tessera sanitaria, cognome, data di nascita
- step 3: tessera sanitaria, nome, data di nascita
- step 4: tessera sanitaria, cognome, nome
- step 5: cognome, nome, data di nascita

Negli archivi dove è presente anche il Codice Fiscale tra le variabili, prima dello step 5 sono stati effettuati i seguenti passi:

- step 1 (CF): codice fiscale, cognome, nome, data di nascita
- step 2 (CF): codice fiscale, cognome, data di nascita
- step 3 (CF): codice fiscale, nome, data di nascita

step 4 (CF): codice fiscale, cognome, nome

Inoltre, fino al 2007, per gli stessi sono stati aggiunti ulteriori quattro passi di chiave, corrispondenti agli step n (CF) sopra riportati, ma utilizzando il Codice Fiscale ricalcolato dell'Anagrafe Sanitaria.

Questa procedura ha permesso di identificare soltanto poche unità aggiuntive di record e non è stata ripetuta per l'anno 2008.

Nell'archivio di mortalità non è presente la Tessera Sanitaria, quindi nei primi 4 step è sostituita con il Codice Fiscale.

Infine, l'incrocio dei farmaci con l'Anagrafe avviene mediante 2 soli step, essendo TS o CF l'unica informazione anagrafica contenuta in questo archivio.

Volendo dare una visione d'insieme, nella seguente tabella sono evidenziate le caselle corrispondenti ai passi di chiave utilizzati nel record linkage dello specifico archivio (non sono riportati i due passi semplici per le Prescrizioni farmaceutiche):

**Tab. 3.10:** Passi di chiave utilizzati nel record linkage, per archivio

Passi di chiave	SDO	mortalità	MP	AP	ET
TS C N DN					
T C DN					
T C N					
T N DN					
CF C N DN					
CF C DN					
CF C N					
CF N DN					
C N DN					
nuovo_CF C N DN					
nuovo_CF C DN					
nuovo_CF C N					
nuovo_CF N DN					

**Legenda:**

T: Tessera Sanitaria

C: Cognome

N: Nome

DN: Data di Nascita

CF: Codice Fiscale

nuovo\_CF: Codice Fiscale ricalcolato

Data l'esiguità dei record agganciati mediante i passi di chiave contenenti il CF ricalcolato, nelle tabelle successive gli ultimi 4 passi sono stati raggruppati assieme.

### 3.3 - Esecuzione di procedure di record linkage semi-deterministico a cascata

E' stata elaborata una procedura per effettuare il record linkage in Access, con l'obiettivo di massimizzare standardizzazione (replicabilità) e sicurezza.

La procedura è stata elaborata sostanzialmente de novo.

Sono state formulate una serie di query eseguite a cascata; la logica di ogni ciclo di query è selezionare i soggetti che linkano per il passo di chiave, sottrarli dall'archivio con il quale si è effettuato il linkage e creare il nuovo archivio dopo la sottrazione, oltre ad una tabella contenente i record linkati. Al termine dello step "n" si otterranno dunque due tabelle:

- tab\_stepn: contiene i record linkati nello step n
- tab\_dopo\_stepn: contiene i record dell'archivio che non sono ancora stati linkati fino a questo punto (è la nuova tabella che verrà incrociata con l'anagrafe allo step n+1)

Nella query di incrocio viene aggiunta ad ogni record la variabile Tessera Sanitaria, ottenuta dall'Anagrafe.

Dall'unione di tutte le tabelle tab\_stepn si ottiene il dataset con tutti i record agganciati; in questo dataset si eliminano le variabili anagrafiche dell'archivio, poiché le informazioni anagrafiche utilizzate in seguito saranno esclusivamente derivate dall'Anagrafe Sanitaria.

L'incrocio delle Prescrizioni farmaceutiche con l'Anagrafe richiede di lavorare su diversi file Access, dato l'elevato numero di record di questo archivio.

### 3.4 - Valutazione delle percentuali di linkage

Si riportano di seguito le tabelle con il numero assoluto di record linkati per singolo passo di chiave, il numero degli appaiati cumulati nonché dei record non agganciati, con le relative percentuali.

#### Mortalità

**Tab. 3.11:** Percentuali di incrocio per singolo passo di chiave e totale - mortalità

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
mortalità	1996-2008	K0	28663			
		CFCNDN	14118	14118	49,26	49,3
		CFCDN	203	14321	0,71	50,0
		CFCN	45	14366	0,16	50,1
		CFNDN	207	14573	0,72	50,8
		CNDN	9162	23735	31,96	82,8
		nuovo_cf*	102	23837	0,36	83,2

\* incrocio non effettuato per 2008

<b>non appaiati</b>	<b>4826</b>	<b>16,8</b>
---------------------	-------------	-------------

**Tab. 3.12:** Percentuali di record appaiati per anno - mortalità

Anno	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
% appaiati cumulati	75.3	73	76.2	80.6	79.2	82.4	84	84.2	85.9	86	90.6	91.4	90.8
% appaiati cumul. ASL					92.3	95.6	95	93.7	94.7	95	99.6	99.1	99.4

La quota di non appaiati è molto elevata.

Dal 2000 si è calcolata la percentuale di soggetti assistiti dall'ASL 18 non agganciati con l'Anagrafe: questa quota è in diminuzione, passando dal massimo di 7,7% del 2000 allo 0,6% del 2008 ed attestandosi a meno dell'1% anche nei 2 anni precedenti.

Schede di dimissione ospedaliera (SDO)

**Tab. 3.13:** Percentuali di incrocio per singolo passo di chiave e totale - SDO

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
SDO	1997-2008	K0	446404			
		TCNDN	285787	285787	64,02	64,0
		TCDN	8493	294280	1,90	65,9
		TCN	2325	296605	0,52	66,4
		TNDN	1015	297620	0,23	66,7
		CFCNDN	12492	310112	2,80	69,5
		CFCDN	289	310401	0,06	69,5
		CFCN	18	310419	0,00	69,5
		CFNDN	136	310555	0,03	69,6
		CNDN	4803	315358	1,08	70,6
		nuovo_cf*	53	315411	0,01	70,7

\* incrocio non effettuato per 2008

non appaiati	130993	29,3
di cui dell'ASL	3561	2,7

**Tab. 3.14:** Percentuali di record appaiati per anno - SDO

Anno	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
% appaiati cumulati	72.1	72.8	72.5	71.4	71.5	71.4	71.3	70.7	66.7	66.8	68.3	66.6
% appaiati cumulati ASL				98.5	98.4	98.4	98.5	98.4	98.6	98.8	98.6	98.4

L'archivio SDO contiene tutti i ricoveri effettuati dai presidi ospedalieri dell'ASL 18 e non soltanto quelli riferiti ad assistiti della stessa. La quota di assistiti non appaiati sul totale dei record è pari allo 0,8%.

Mobilità passiva delle SDO

Tab. 3.15: Percentuali di incrocio per singolo passo di chiave e totale – Mobilità passiva intra-regionale

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
MP intra	2000-2007	K0	<b>40172</b>			
		TCNDN	28577	28577	71,14	71,1
		TCDN	1338	29915	3,33	74,5
		TCN	228	30143	0,57	75,0
		TNDN	256	30399	0,64	75,7
		CFCNDN	7540	37939	18,77	94,4
		CFCDN	155	38094	0,39	94,8
		CFCN	10	38104	0,02	94,9
		CFNDN	35	38139	0,09	94,9
		CNDN	752	38891	1,87	96,8
		nuovo_cf	7	<b>38898</b>	0,02	96,8

non appaiati 1274 3,2

Tab. 3.16: Percentuali di incrocio per singolo passo di chiave e totale – Mobilità passiva extra-regionale

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
MP extra	2000-2007	K0	<b>33808</b>			
		TCNDN	846	846	2,50	2,5
		TCDN	56	902	0,17	2,7
		TCN	11	913	0,03	2,7
		TNDN	11	924	0,03	2,7
		CFCNDN	22552	23476	66,71	69,4
		CFCDN	544	24020	1,61	71,0
		CFCN	2	24022	0,01	71,1
		CFNDN	307	24329	0,91	72,0
		CNDN	2664	26993	7,88	79,8
		nuovo_cf	16	<b>27009</b>	0,05	79,9

non appaiati 6799 20,1  
non appaiati dell'ASL 6719 19,9

**Tab. 3.17:** Percentuali di record appaiati per anno – mobilità passiva delle SDO

Anno	2000	2001	2002	2003	2004	2005	2006	2007	2008
% appaiati cumulati INTRA	96.6	97.5	96.5	97.1	97.5	97.1	97	95.5	86.5
% appaiati cumulati EXTRA	64.8	85.3	70.9	88.3	82.5	88.4	86.8	71.8	

Per quanto riguarda la mobilità passiva extra-regionale, la percentuale di assistiti non appaiati è molto elevata. Si suppone sia dovuta a difformità nell'archiviazione dei dati in altre realtà regionali.

### Prescrizioni farmaceutiche

**Tab. 3.18:** Percentuali di incrocio per singolo passo di chiave e totale – Prescrizioni farmaceutiche

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
farmaci	97-08	K0	16443052			
		TS	15731134	15731134	95,67	95,7
		CF	710	15731844	0,00	95,7

**non appaiati      711208      4,3**

**Tab. 3.19:** Percentuali di record appaiati per anno – Prescrizioni farmaceutiche

Anno	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
% appaiati cumulati	92.5	87.9	93.4	94.1	95.8	95.4	96.9	97.2	97	97.6	97.8	98.1

Dall'anno 2003 la percentuale di record non appaiati è inferiore al 3%.

### Esenzioni Ticket (ET)

**Tab. 3.20:** Percentuali di incrocio per singolo passo di chiave e totale – Esenzioni Ticket

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
esenzioni	1996-2007	K0	131384			
		CFCNDN	131035	131035	99,73	99,7
		CFCDN	146	131181	0,11	99,8
		CFCN	16	131197	0,01	99,9
		CFNDN	15	131212	0,01	99,9
		CNDN	6	131218	0,00	99,9
		nuovo_cf	19	131237	0,01	99,9

**non appaiati      166      0,1**

L'archivio delle ET è gestito in modo analogo all'Anagrafe; la quota di linkage è pressoché totale.

**Tab. 3.21:** Percentuali di record appaiati per anno – Esenzioni Ticket

Anno	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
% appaiati cumulati	100	100	99.9	99.9	100	99.9	99.9	99.9	99.9	99.8	99.6	99.7	97.3

## Fase 4: produzione di stime di frequenza

Tempo di realizzazione: settembre – novembre 2009

### 4.1 - Applicazione agli archivi di algoritmi forniti

Gli algoritmi applicati alla popolazione dell'ASL 18 corrispondono a quelli presentati nel Supplemento della Rivista "Epidemiologia e Prevenzione", Anno 32 (3), maggio-giugno 2008.

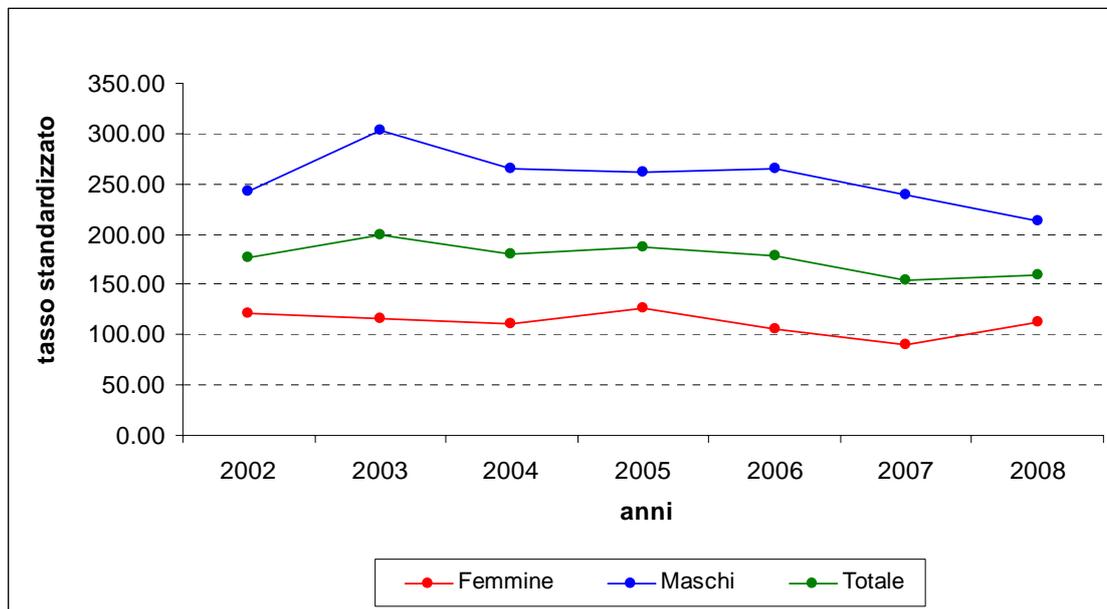
## Risultati

IMA

**Tab.4.1-IMA:** Incidenza grezza e standardizzata dell'IMA (\*100 000 abitanti)

anno	sexso	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2002	F	153	88174	173.52	<b>120.83</b>
2002	M	193	81932	235.56	<b>242.20</b>
2002	totale	346	170106	203.40	<b>177.30</b>
2003	F	148	88193	167.81	<b>115.59</b>
2003	M	244	82130	297.09	<b>303.02</b>
2003	totale	392	170323	230.15	<b>199.05</b>
2004	F	148	88789	166.69	<b>111.05</b>
2004	M	217	82639	262.59	<b>265.75</b>
2004	totale	365	171428	212.92	<b>180.34</b>
2005	F	167	89052	187.53	<b>125.93</b>
2005	M	211	83099	253.91	<b>261.60</b>
2005	totale	378	172151	219.57	<b>186.92</b>
2006	F	139	89084	156.03	<b>106.46</b>
2006	M	224	83336	268.79	<b>264.89</b>
2006	totale	363	172420	210.53	<b>178.71</b>
2007	F	123	89235	137.84	<b>90.60</b>
2007	M	201	83502	240.71	<b>238.25</b>
2007	totale	324	172737	187.57	<b>153.68</b>
2008	F	154	89893	171.31	<b>112.08</b>
2008	M	187	84152	222.22	<b>213.34</b>
2008	totale	341	174045	195.93	<b>159.69</b>

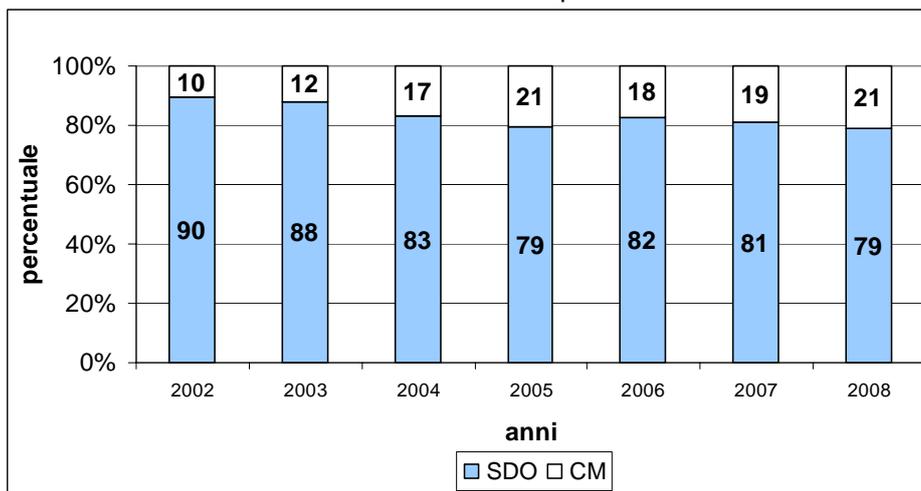
**Graf. 4.1-IMA:** Incidenza standardizzata dell'IMA (\*100 000 abitanti)



**Tab. 4.2-IMA:** Distribuzione casi per classi d'età e anno

Classe età	Anno							Totale
	2002	2003	2004	2005	2006	2007	2008	
<b>0-14</b>	0	0	1	1	0	0	0	2
	0	0	0.27	0.26	0	0	0	
<b>15-34</b>	3	3	0	1	5	0	0	12
	0.87	0.77	0	0.26	1.38	0	0	
<b>35-54</b>	45	36	53	37	49	37	38	295
	13.01	9.18	14.52	9.79	13.5	11.42	11.14	
<b>55-64</b>	45	66	33	49	55	47	48	343
	13.01	16.84	9.04	12.96	15.15	14.51	14.08	
<b>65-74</b>	82	107	84	72	79	70	81	575
	23.7	27.3	23.01	19.05	21.76	21.6	23.75	
<b>75+</b>	171	180	194	218	175	170	174	1282
	49.42	45.92	53.15	57.67	48.21	52.47	51.03	
<b>Totale</b>	346	392	365	378	363	324	341	2509

**Graf. 4.2-IMA:** Contributo esclusivo delle fonti per anno di stima

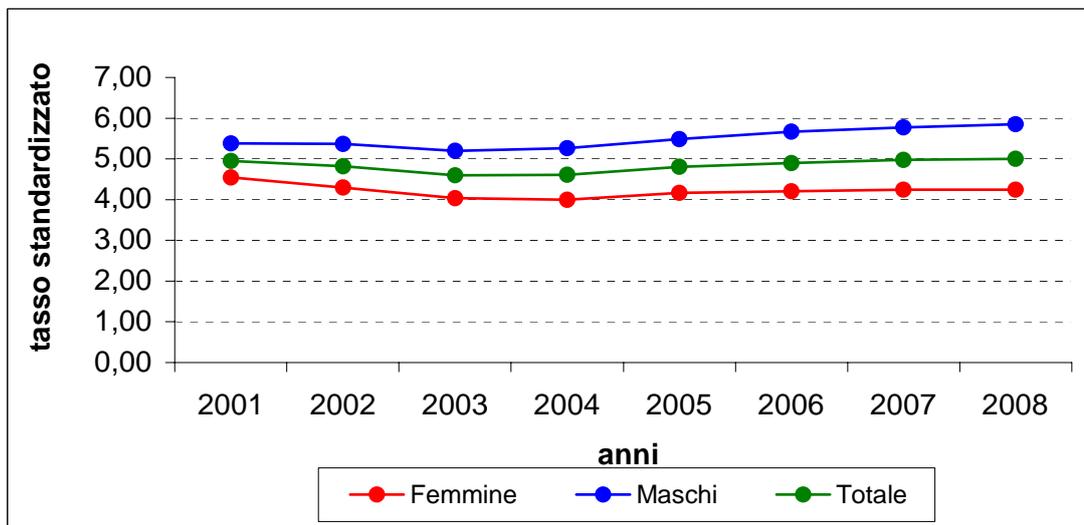


Diabete

**Tab. 4.1-diabete:** Prevalenza grezza e standardizzata del diabete (\*100 abitanti)

anno	sessot	numeratore	denominatore	tasso grezzo	tasso stand
2001	F	4876	88174	5,53	4,54
2001	M	4433	81932	5,41	5,39
2001	totale	9309	170106	5,47	4,95
2002	F	4682	88174	5,31	4,30
2002	M	4414	81932	5,39	5,37
2002	totale	9096	170106	5,35	4,82
2003	F	4494	88193	5,10	4,04
2003	M	4326	82130	5,27	5,20
2003	totale	8820	170323	5,18	4,60
2004	F	4499	88789	5,07	4,00
2004	M	4428	82639	5,36	5,26
2004	totale	8927	171428	5,21	4,60
2005	F	4726	89052	5,31	4,17
2005	M	4683	83099	5,64	5,49
2005	totale	9409	172151	5,47	4,80
2006	F	4788	89084	5,37	4,21
2006	M	4877	83336	5,85	5,66
2006	totale	9665	172420	5,61	4,90
2007	F	4846	89235	5,43	4,24
2007	M	5038	83502	6,03	5,78
2007	totale	9884	172737	5,72	4,97
2008	F	4891	89893	5,44	4,24
2008	M	5153	84152	6,12	5,86
2008	totale	10044	174045	5,77	5,01

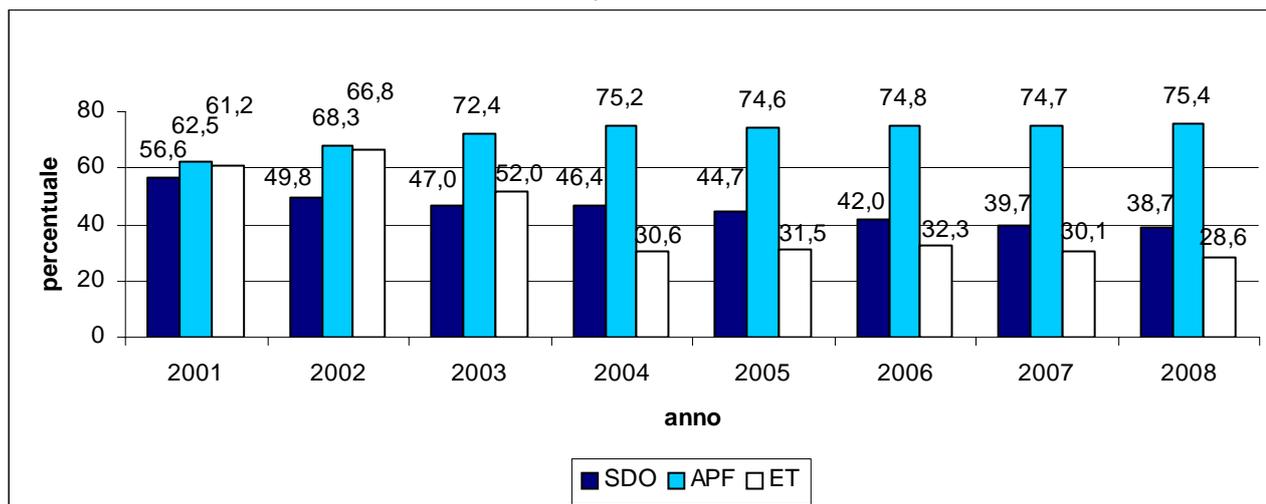
**Graf. 4.1-diabete:** Prevalenza standardizzata del diabete (\*100 abitanti)



**Tab. 4.2-diabete:** Distribuzione casi per classi d'età e anno di stima

Classe età	Anno								Totale
	2001	2002	2003	2004	2005	2006	2007	2008	
<b>0-14</b>	77 0.83	43 0.47	20 0.23	22 0.25	28 0.3	25 0.26	26 0.26	25 0.25	266
<b>15-34</b>	364 3.91	266 2.92	183 2.07	185 2.07	185 1.97	175 1.81	209 2.11	208 2.07	1775
<b>35-54</b>	1512 16.24	1377 15.1	1271 14.41	1234 13.8	1278 13.58	1272 13.2	1311 13.26	1281 12.75	10536
<b>55-64</b>	1850 19.87	1903 20.9	1817 20.6	1834 20.5	1970 20.94	2065 21.4	2132 21.57	2183 21.73	15754
<b>65-74</b>	2765 29.7	2717 29.9	2681 30.4	2745 30.8	2797 29.73	2847 29.5	2858 28.92	2919 29.06	22335
<b>75+</b>	2741 29.44	2790 30.7	2848 32.29	2907 32.6	3151 33.49	3281 34	3348 33.87	3428 34.13	24497
<b>Totale</b>	9309	9096	8820	8927	9409	9665	9884	10044	75163

**Graf. 4.2-diabete:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3-diabete:** Contributo esclusivo delle fonti, anni di stima 2004-2008

Fonte	2004		2005		2006		2007		2008	
	Freq.	%								
<b>SDO</b>	1188	13.3	1241	13.2	1169	12.1	1179	11.9	1174	11.7
<b>APF</b>	2857	32.0	3078	32.7	3306	34.2	3681	37.2	3940	39.2
<b>ET</b>	858	9.6	958	10.2	1067	11.0	1134	11.5	1115	11.1
<b>più di una fonte</b>	4024	45.1	4132	43.9	4123	42.7	3890	39.4	3815	38.0
<b>totale</b>	8927	100	9409	100	9665	100	9884	100	10044	100.0

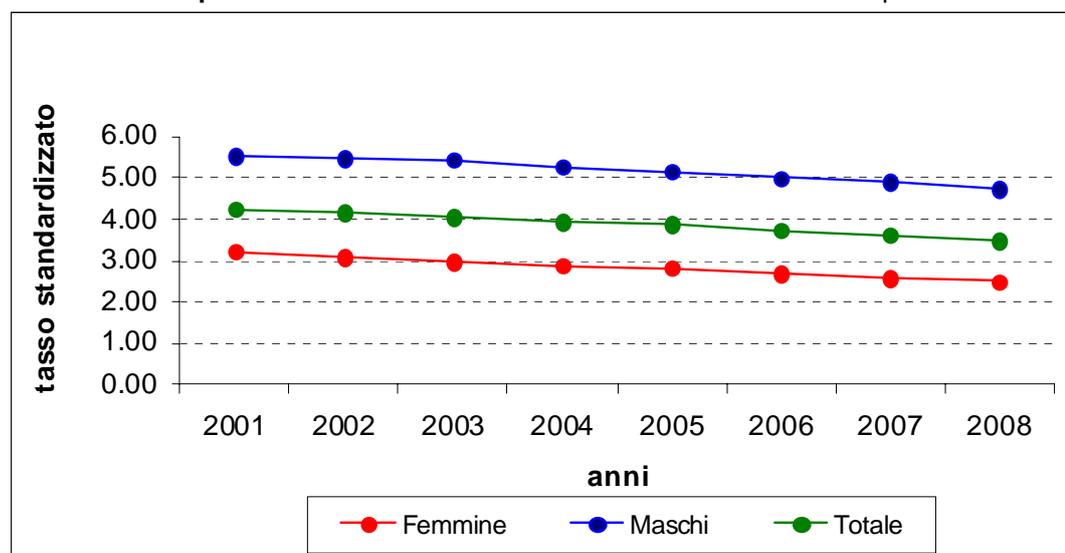
## Cardiopatía ischemica

**Tab. 4.1-cardiopatía ischemica:** Prevalenza grezza e standardizzata della cardiopatía ischemica (\*100 abitanti)

anno	sessot	numeratore	denominatore	tasso grezzo	tasso grezzo	tasso grezzo	tasso stand
2001	F	4028	88174	4,57	3,22		
2001	M	4331	81932	5,29	5,52		
2001	totale	8359	170106	4,91	4,25		
2002	F	3933	88174	4,46	3,10		
2002	M	4266	81932	5,21	5,48		
2002	totale	8199	170106	4,82	4,16		
2003	F	3860	88193	4,38	2,97		
2003	M	4286	82130	5,22	5,43		
2003	totale	8146	170323	4,78	4,06		
2004	F	3794	88789	4,27	2,88		
2004	M	4218	82639	5,10	5,26		
2004	totale	8012	171428	4,67	3,94		
2005	F	3793	89052	4,26	2,83		
2005	M	4181	83099	5,03	5,16		

2005	totale	7974	172151	4,63	<b>3,86</b>
2006	F	3707	89084	4,16	<b>2,70</b>
2006	M	4133	83336	4,96	<b>5,00</b>
2006	totale	7840	172420	4,55	<b>3,73</b>
2007	F	3608	89235	4,04	<b>2,59</b>
2007	M	4121	83502	4,94	<b>4,91</b>
2007	totale	7729	172737	4,47	<b>3,62</b>
2008	F	3550	89893	3,95	<b>2,49</b>
2008	M	4035	84152	4,79	<b>4,73</b>
2008	totale	7585	174045	4,36	<b>3,50</b>

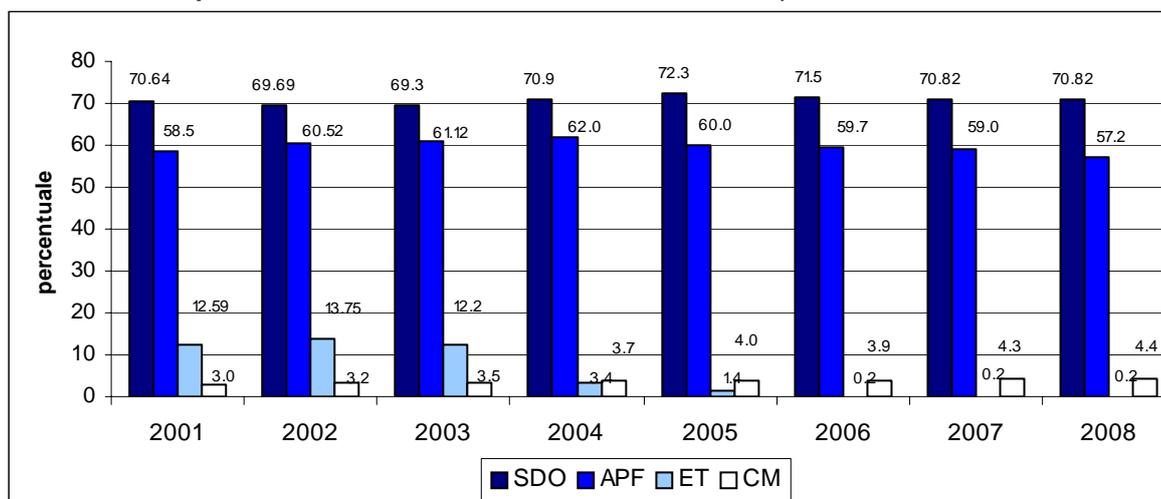
Graf. 4.1-cardiopatia ischemica: Prevalenza standardizzata della cardiopatia ischemica (\*100 abitanti)



Tab. 4.2-cardiopatia ischemica: Distribuzione casi per classi d'età e anno di stima

Classe età	Anno								Totale
	2001	2002	2003	2004	2005	2006	2007	2008	
<b>0-14</b>	13 0.16	9 0.11	8 0.10	10 0.12	9 0.11	8 0.10	7 0.09	4 0.05	68
<b>15-34</b>	63 0.75	24 0.29	10 0.12	7 0.09	9 0.11	20 0.26	20 0.26	15 0.20	168
<b>35-54</b>	601 7.19	531 6.48	450 5.52	446 5.57	397 4.98	411 5.24	419 5.42	428 5.64	3683
<b>55-64</b>	993 11.88	961 11.72	935 11.48	893 11.15	895 11.22	901 11.49	916 11.85	898 11.84	7392
<b>65-74</b>	2332 27.90	2201 26.84	2157 26.48	2056 25.66	1947 24.42	1845 23.53	1757 22.73	1687 22.24	15982
<b>75+</b>	4357 52.12	4473 54.56	4586 56.30	4600 57.41	4717 59.15	4655 59.38	4610 59.65	4553 60.03	36551
<b>Totale</b>	8359	8199	8146	8012	7974	7840	7729	7585	63844

**Graf. 4.2-cardiopatia ischemica: Contributo assoluto delle fonti per anno di stima**



**Tab. 4.3- cardiopatia ischemica: Contributo esclusivo delle fonti, anni di stima 2004-2008**

Fonte	2004		2005		2006		2007		2008	
	Freq.	%								
<b>SDO</b>	2704	33.8	2913	36.5	2938	37.5	2933	38.0	2996	39.5
<b>APF</b>	2171	27.1	2075	26.0	2120	27.0	2128	27.5	2092	27.6
<b>ET</b>	46	0.6	14	0.2	2	0.0	2	0.0	2	0.0
<b>CM</b>	73	0.9	97	1.2	97	1.2	101	1.3	106	1.4
<b>più di una fonte</b>	3018	37.7	2875	36.1	2683	34.2	2565	33.2	2389	31.5
<b>totale</b>	8012	100	7974	100	7840	100	7729	100	7585	100.0

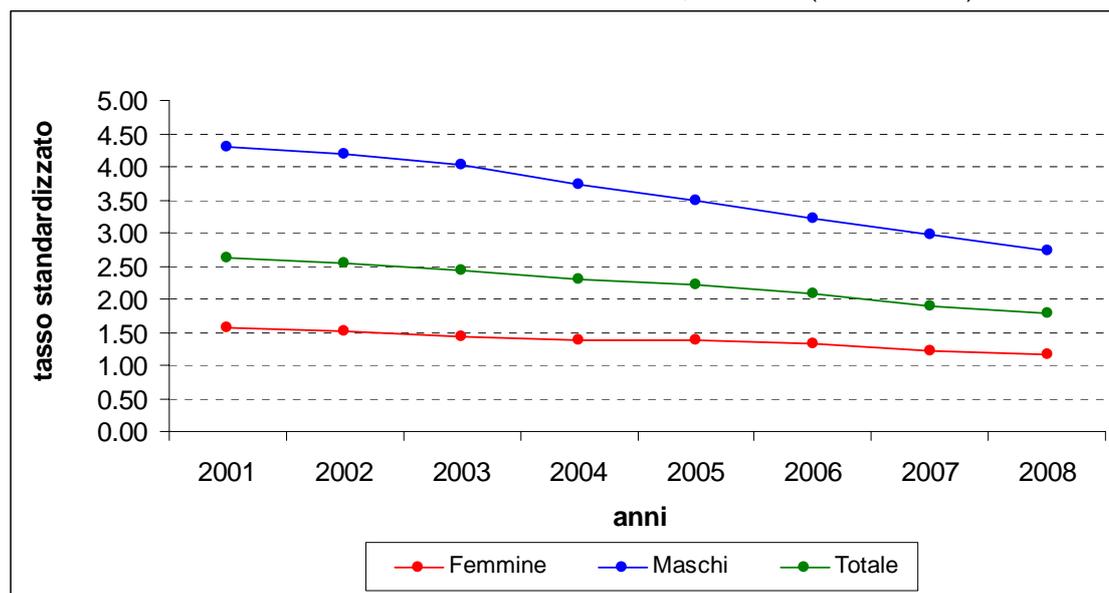
**BPCO**

**Tab. 4.1-BPCO: Prevalenza grezza e standardizzata della BPCO, 35+ anni (\*100 abitanti)**

anno	sess	numeratore	denominatore	tasso grezzo	tasso stand
anno	sess	tasso grezzo	tasso grezzo	tasso grezzo	tasso stand
2001	F	1102	58597	1,88	<b>1,56</b>
2001	M	1957	50744	3,86	<b>4,29</b>
2001	totale	3059	109341	2,80	<b>2,63</b>
2002	F	1086	58597	1,85	<b>1,51</b>
2002	M	1890	50744	3,72	<b>4,18</b>
2002	totale	2976	109341	2,72	<b>2,55</b>
2003	F	1066	59099	1,80	<b>1,44</b>
2003	M	1846	51408	3,59	<b>4,03</b>
2003	totale	2912	110507	2,64	<b>2,44</b>
2004	F	1052	59836	1,76	<b>1,39</b>
2004	M	1741	52049	3,34	<b>3,74</b>
2004	totale	2793	111885	2,50	<b>2,31</b>
2005	F	1067	60364	1,77	<b>1,39</b>
2005	M	1656	52744	3,14	<b>3,48</b>

2005	totale	2723	113108	2,41	<b>2,22</b>
2006	F	1048	60637	1,73	<b>1,33</b>
2006	M	1560	53276	2,93	<b>3,23</b>
2006	totale	2608	113913	2,29	<b>2,08</b>
2007	F	976	61085	1,60	<b>1,21</b>
2007	M	1460	53795	2,71	<b>2,96</b>
2007	totale	2436	114880	2,12	<b>1,90</b>
2008	F	950	61612	1,54	<b>1,16</b>
2008	M	1360	54585	2,49	<b>2,73</b>
2008	totale	2310	116197	1,99	<b>1,77</b>

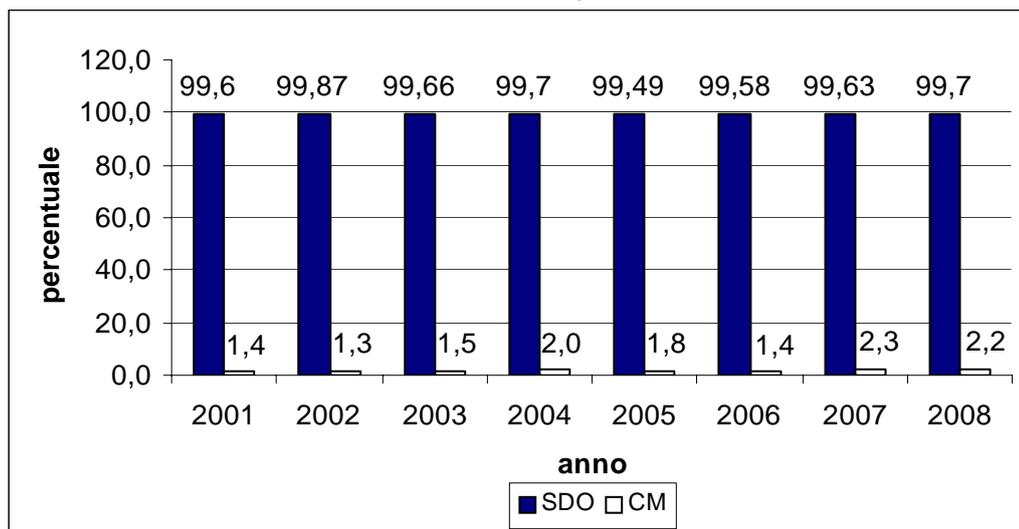
**Graf. 4.1-BPCO:** Prevalenza standardizzata della BPCO, 35+ anni (\*100 abitanti)



**Tab. 4.2-BPCO:** Distribuzione casi per classi d'età e anno di stima

Classe età	Anno								Totale
	2001	2002	2003	2004	2005	2006	2007	2008	
<b>35-54</b>	270 8.83	223 7.49	179 6.15	173 6.19	150 5.51	130 4.98	133 5.46	120 5.19	1378
<b>55-64</b>	347 11.3	315 10.58	283 9.72	262 9.38	258 9.47	250 9.59	220 9.03	202 8.74	2137
<b>65-74</b>	955 31.2	877 29.47	819 28.1	719 25.7	657 24.1	565 21.7	528 21.7	501 21.69	5621
<b>75+</b>	1487 48.6	1561 52.45	1631 56	1639 58.7	1658 60.9	1663 63.8	1555 63.8	1487 64.37	12681
<b>Totale</b>	3059	2976	2912	2793	2723	2608	2436	2310	21817

**Graf. 4.2-BPCO:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- BPCO:** Contributo esclusivo delle fonti, anni di stima 2004-2008

Fonte	2004		2005		2006		2007		2008	
	Freq.	%								
<b>SDO</b>	2736	98.0	2675	98.2	2571	98.6	2381	97.7	2260	97.8
<b>CM</b>	9	0.3	14	0.5	11	0.4	9	0.4	6	0.3
<b>più di una fonte</b>	48	1.7	34	1.3	26	1.0	46	1.9	44	1.9
<b>totale</b>	2793	100	2723	100	2608	100	2436	100	2310	100

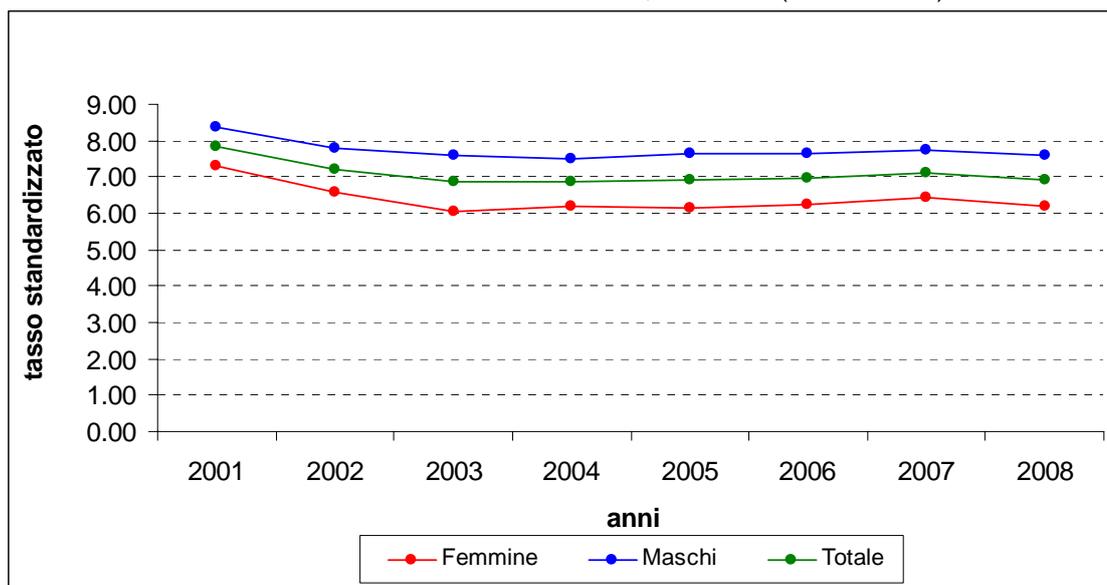
## Asma

**Tab. 4.1-asma:** Prevalenza grezza e standardizzata dell'asma, 0-34 anni (\*100 abitanti)

anno	sess	numeratore	denominatore	tasso grezzo	tasso stand
		tasso grezzo	tasso grezzo	tasso grezzo	tasso stand
2001	F	2016	29577	6,82	<b>7,29</b>
2001	M	2417	31188	7,75	<b>8,36</b>
2001	totale	4433	60765	7,30	<b>7,84</b>
2002	F	1812	29577	6,13	<b>6,58</b>
2002	M	2243	31188	7,19	<b>7,79</b>
2002	totale	4055	60765	6,67	<b>7,20</b>
2003	F	1665	29094	5,72	<b>6,07</b>
2003	M	2180	30722	7,10	<b>7,61</b>
2003	totale	3845	59816	6,43	<b>6,86</b>
2004	F	1711	28953	5,91	<b>6,18</b>
2004	M	2152	30590	7,03	<b>7,50</b>
2004	totale	3863	59543	6,49	<b>6,85</b>
2005	F	1703	28688	5,94	<b>6,13</b>
2005	M	2208	30355	7,27	<b>7,65</b>
2005	totale	3911	59043	6,62	<b>6,91</b>

2006	F	1734	28447	6,10	<b>6,22</b>
2006	M	2224	30060	7,40	<b>7,65</b>
2006	totale	3958	58507	6,77	<b>6,95</b>
2007	F	1791	28150	6,36	<b>6,43</b>
2007	M	2272	29707	7,65	<b>7,76</b>
2007	totale	4063	57857	7,02	<b>7,12</b>
2008	F	1758	28281	6,22	<b>6,21</b>
2008	M	2248	29567	7,60	<b>7,61</b>
2008	totale	4006	57848	6,93	<b>6,93</b>

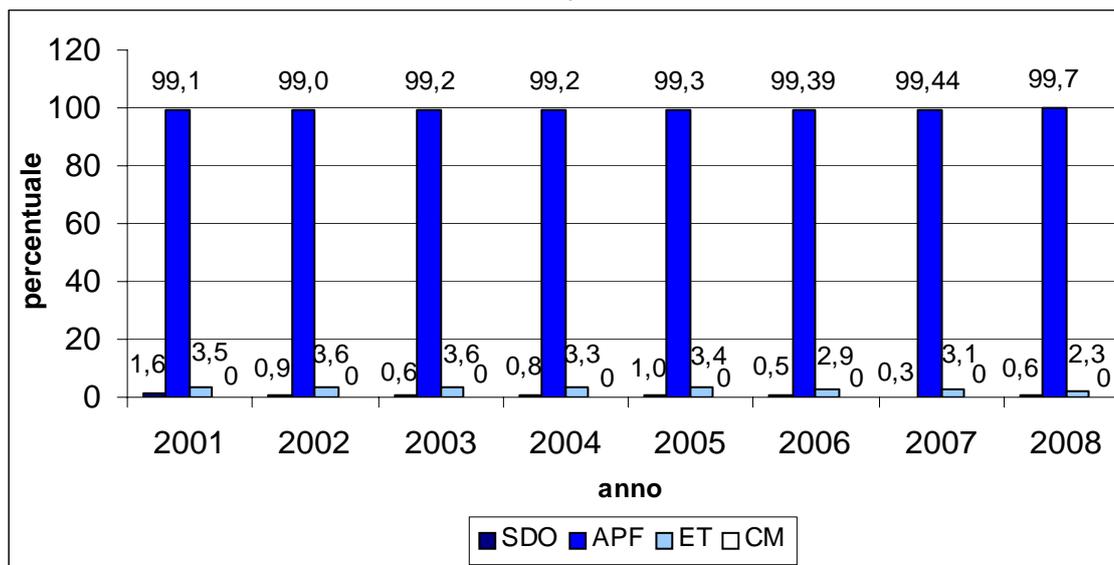
**Graf. 4.1-asma:** Prevalenza standardizzata dell'asma, 0-34 anni (\*100 abitanti)



**Tab. 4.2-asma:** Distribuzione casi per classi d'età e anno di stima

Classe età	Anno								Totale
	2001	2002	2003	2004	2005	2006	2007	2008	
0-14	2898	2744	2529	2497	2561	2604	2571	2629	21033
	65.37	67.67	65.77	64.64	65.48	65.79	62.98	65.63	
15-34	1535	1311	1316	1366	1350	1354	1511	1377	11120
	34.63	32.33	34.23	35.36	34.52	34.21	37.02	34.37	
<b>Totale</b>	4433	4055	3845	3863	3911	3958	4082	4006	32153

**Graf. 4.2-asma:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- asma:** Contributo esclusivo delle fonti, anni di stima 2004-2008

Fonte	2004		2005		2006		2007		2008	
	Freq.	%								
<b>SDO</b>	8	0,2	11	0,3	2	0,1	0	0	2	0,1
<b>APF</b>	3708	96,0	3748	95,8	3829	96,7	3947	96,7	3892	97,2
<b>ET</b>	22	0,6	18	0,5	22	0,6	22	0,5	10	0,3
<b>CM</b>	0	0	0	0	0	0	0	0	0	0
<b>più di una fonte</b>	125	3,2	134	3,4	105	2,7	113	2,8	102	2,6
<b>totale</b>	3863	100	3911	100	3958	100	4082	100	4006	100

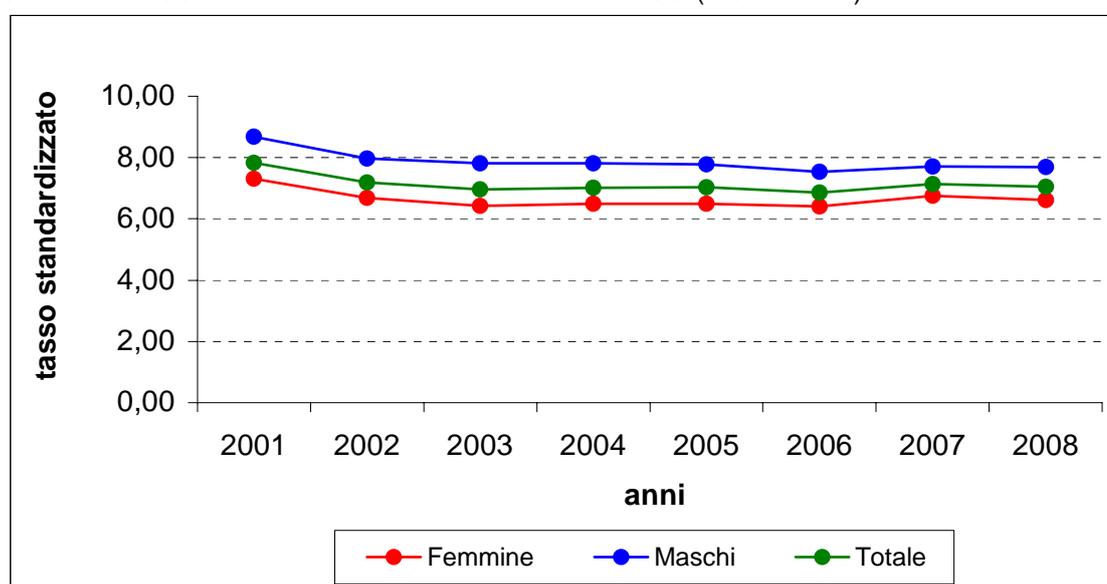
MPCO

**Tab. 4.1-MPCO:** Prevalenza grezza e standardizzata delle MPCO (\*100 abitanti)

anno	sessot	numeratore	denominatore	tasso grezzo	tasso stand
2001	F	6533	88174	7,41	<b>7,31</b>
2001	M	6644	81932	8,11	<b>8,67</b>
2001	totale	13177	170106	7,75	<b>7,82</b>
2002	F	6047	88174	6,86	<b>6,68</b>
2002	M	6091	81932	7,43	<b>7,97</b>
2002	totale	12138	170106	7,14	<b>7,19</b>
2003	F	5851	88193	6,63	<b>6,42</b>
2003	M	6017	82130	7,33	<b>7,82</b>
2003	totale	11868	170323	6,97	<b>6,97</b>
2004	F	6008	88789	6,77	<b>6,49</b>
2004	M	6080	82639	7,36	<b>7,81</b>
2004	totale	12088	171428	7,05	<b>7,01</b>
2005	F	6060	89052	6,81	<b>6,50</b>

2005	M	6144	83099	7,39	<b>7,78</b>
2005	totale	12204	172151	7,09	<b>7,03</b>
2006	F	5938	89084	6,67	<b>6,40</b>
2006	M	5985	83336	7,18	<b>7,53</b>
2006	totale	11923	172420	6,92	<b>6,86</b>
2007	F	6296	89235	7,06	<b>6,76</b>
2007	M	6221	83502	7,45	<b>7,71</b>
2007	totale	12517	172737	7,25	<b>7,13</b>
2008	F	6267	89893	6,97	<b>6,61</b>
2008	M	6293	84152	7,48	<b>7,69</b>
2008	totale	12560	174045	7,22	<b>7,05</b>

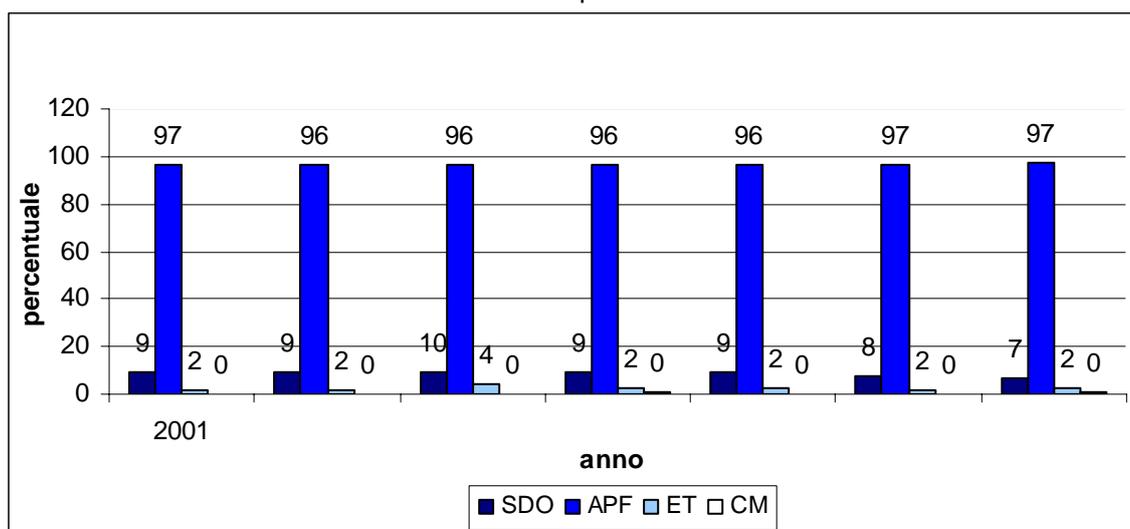
Graf. 4.1-MPCO: Prevalenza standardizzata delle MPCO (\*100 abitanti)



Tab. 4.2-MPCO: Distribuzione casi per classi d'età e anno di stima

Classe età	Anno								Totale
	2001	2002	2003	2004	2005	2006	2007	2008	
0-14	2932	2765	2552	2512	2571	2617	2578	2633	21160
	22.25	22.78	21.5	20.78	21.07	21.95	20.54	20.96	
15-34	1669	1386	1370	1418	1400	1393	1559	1399	11594
	12.67	11.42	11.54	11.73	11.47	11.68	12.42	11.14	
35-54	2184	1951	2004	1955	2037	2110	2346	2404	16991
	16.57	16.07	16.89	16.17	16.69	17.7	18.69	19.14	
55-64	1381	1245	1245	1313	1339	1260	1366	1415	10564
	10.48	10.26	10.49	10.86	10.97	10.57	10.88	11.27	
65-74	2390	2147	2037	2083	2028	1836	1897	1893	16311
	18.14	17.69	17.16	17.23	16.62	15.4	15.11	15.07	
75+	2621	2644	2660	2807	2829	2707	2806	2816	21890
	19.89	21.78	22.41	23.22	23.18	22.7	22.36	22.42	
<b>Totale</b>	13177	12138	11868	12088	12204	11923	12552	12560	98510

**Graf. 4.2-MPCO:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- MPCO:** Contributo esclusivo delle fonti, anni di stima 2004-2008

	2004		2005		2006		2007		2008	
	Freq.	%								
<b>SDO</b>	365	3.0	413	3.4	320	2.7	274	2.2	306	2.4
<b>APF</b>	10706	88.6	10835	88.8	10800	90.6	11405	90.9	11495	91.5
<b>ET</b>	39	0.3	30	0.3	36	0.3	49	0.4	28	0.2
<b>CM</b>	10	0.1	15	0.1	12	0.1	12	0.1	8	0.1
<b>più di una fonte</b>	968	8.0	911	7.5	755	6.3	812	6.5	723	5.8
<b>totale</b>	12088	100	12204	100	11923	100	12552	100	12560	100

#### 4.2 - Confronto con stime ottenute applicando gli stessi algoritmi in altre realtà italiane

Le altre realtà italiane prese come riferimento per il confronto delle stime di incidenza e prevalenza delle patologie di cui sopra sono quelle presenti nel Supplemento della Rivista "Epidemiologia e Prevenzione" sopra citato. Le stime pubblicate in questo documento sono state prodotte per gli anni 2002, 2003 e 2004. Le seguenti considerazioni si riferiscono a confronti in questo periodo temporale.

La realtà più vicina all'ASL 18 di Rovigo che abbia utilizzato questi algoritmi è l'Azienda Sanitaria 12 di Venezia.

Per quanto riguarda l'**IMA**, le stime di incidenza ottenute a Rovigo sono più elevate rispetto a quelle della popolazione veneziana, in entrambi i sessi. I tassi risultano essere in generale più alti nelle femmine, abbastanza in linea nei maschi con gli altri centri.

Il contributo della fonte Mortalità nella definizione della casistica è minore nell'ASL 18 di Rovigo: a Venezia questa fonte identifica quasi il 20% dei casi di IMA. In generale l'andamento dei contributi a Rovigo diviene in linea con gli altri centri dal 2004.

La prevalenza stimata di **diabete** risulta più elevata a Rovigo, in particolare nelle donne e nel totale, rispetto a Venezia. I tassi sono più elevati anche rispetto agli altri centri.

Nei primi anni di stima il contributo assoluto della fonte ET è molto più elevato a Rovigo rispetto alle altre realtà. Nel corso degli anni la distribuzione dei contributi si allinea a quella delle altre realtà italiane: cresce progressivamente il contributo delle Prescrizioni farmaceutiche, diminuisce quello di SDO ed ET.

Il contributo esclusivo di più fonti e della fonte SDO è minore, risulta in linea quello della fonte APF, è più elevato quello delle ET.

I tassi stimati a Rovigo per la **cardiopatía ischemica** sono in generale più elevati rispetto agli altri centri, sia nella popolazione maschile che femminile.

Rispetto all'anno 2003 degli altri centri, il contributo assoluto della fonte farmaceutica è più elevato, mentre quello delle ET è notevolmente inferiore, che a Rovigo diminuisce progressivamente negli anni di stima.

Rispetto quest'ultima osservazione, anche il contributo esclusivo di questa fonte è molto più basso soprattutto rispetto a Venezia. La distribuzione dei contributi esclusivi è in linea solamente con il centro di Firenze, soprattutto negli ultimi 2 anni di stima.

A Rovigo i tassi di prevalenza di **BPCO** negli uomini sono quasi doppi rispetto a Venezia, mentre sono di poco più elevati nelle donne; l'andamento marcatamente decrescente negli anni porta nel 2008 i tassi in linea con quelli veneziani del periodo 2002-2004.

La distribuzione dei contributi, sia assoluti che esclusivi, è comparabile a quella degli altri centri.

La prevalenza stimata di **asma** è più elevata a Rovigo rispetto che a Venezia e notevolmente più elevata rispetto agli altri centri, sia nella popolazione maschile che femminile.

Analizzando i contributi esclusivi si osserva una più alta percentuale di casi individuati con la sola fonte farmaceutica e una minor quota definita con il contributo di più fonti, rispetto alle altre aree di stima.

Anche le stime di prevalenza di **MPCO** sono più alte a Rovigo, con andamento piuttosto stabile nel periodo. Questo risente verosimilmente della quota di asmatici contenuti in questa categoria di stima.

La distribuzione dei contributi non presenta difformità rispetto agli altri centri di riferimento.

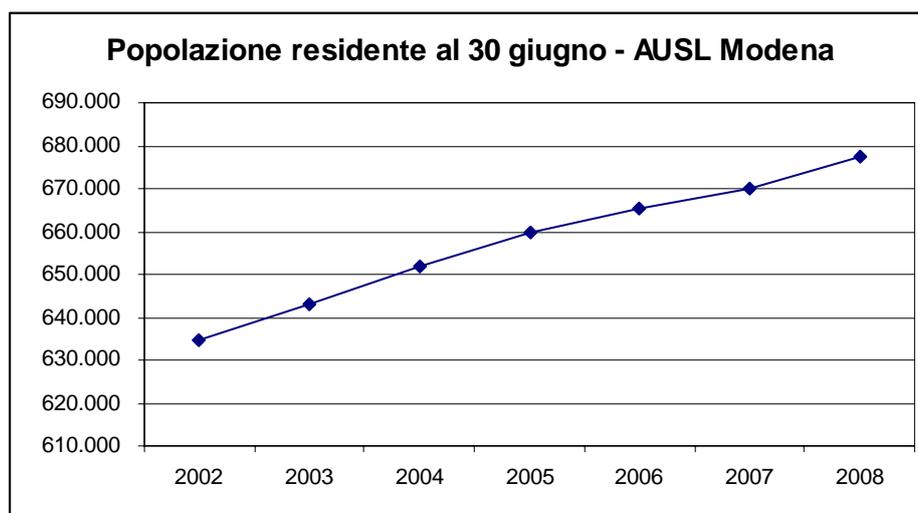
## AUSL Modena

Il territorio dell'Azienda Unità Sanitaria Locale di Modena che coincide con l'intera provincia di Modena comprende 47 Comuni.

La seguente tabella riporta la popolazione residente in questi 47 Comuni al 30 giugno degli anni dal 2002 al 2008. Si evidenzia una tendenza all'aumento della popolazione

Anno	Popolazione
2002	634.713
2003	643.043
2004	651.996
2005	659.925
2006	665.367
2007	670.098
2008	677.672

Fonte: ISTAT ([www.demo.istat.it](http://www.demo.istat.it))



## **Fase 1: sistemazione archivi**

Tempo di realizzazione: 1 Luglio – 30 Settembre 2009

### ***1.1 - Acquisizione degli archivi aziendali richiesti per il progetto***

Gli archivi dell'AUSL di Modena disponibili al 1 gennaio 2009 sono i seguenti:

**Tab. 1.1:** Archivi AUSL Modena richiesti per il progetto; sono specificati primo e ultimo anno disponibile ed il formato dei dati

<b>FONTE</b>	<b>DA (anno)</b>	<b>A (anno)</b>
ANAGRAFE SANITARIA		2009
MORTALITA'	1996	2007
SCHEDE DIMISSIONI OSPEDALIERE (SDO), compresa MOBILITA' PASSIVA	2002	2008
PRESCRIZIONI FARMACEUTICHE	2003	2008
ESENZIONI TICKET		2008

Tutti gli archivi sono stati acquisiti ed organizzati in dataset di Microsoft SQL Server 2005.

### ***1.2 - Organizzazione dei dati in un unico dataset per archivio***

Gli archivi, ad eccezione delle Esenzioni Ticket della Anagrafe sanitaria, sono stati ottenuti in file separati per anno per cui si è provveduto a sistamarli in un unico archivio. In questa fase non sono emerse particolari problematiche.

## **Fase 2: valutazione della qualità dei dati**

Tempo di realizzazione: 1 Luglio – 30 Settembre 2009

### ***2.1 - Distribuzioni di qualità per l'Anagrafe Sanitaria***

La chiave dell'archivio è la variabile numerica RNLIB

L'analisi della qualità del campo RNLIB è avvenuta tramite la ricerca di possibili soggetti duplicati in base a cognome, nome e data nascita (questa è la chiave che ovviamente trova più duplicato e che incorpora anche i risultati delle altre chiavi). Questa ricerca ha trovato 13027 record duplicati pari al 0,8% sul totale dei record presenti nella banca dati; Il 92% di questi è già identificato dal sistema come errore ed è stato fuso ad altro rnlb.

Si è ritenuto di valutare l'attendibilità del campo "Posizione assistito", per il quale si riportano legenda e distribuzione:

**Tab. 2.1:** Legenda codici di sospensione

<b>Codice</b>	<b>Descrizione</b>
<b>null</b>	ASSISTITO ATTIVO
<b>1</b>	SCAD.DEROGA FUORI ASL
<b>3</b>	SCAD.ASS.DIMORA FUORI ASL
<b>4</b>	RICUSAZIONE DA PARTE DEL MEDICO
<b>5</b>	SERVIZIO MILITARE
<b>6</b>	CANCELLAZIONE D'UFFICIO
<b>7</b>	2CAD.MODELLO ESTERO
<b>8</b>	SUPERAMENTO ETA' PEDIATRICA
<b>9</b>	SOGGIORNO TEMPORANEO ALL' ESTERO
<b>A</b>	ASSISTENZA TRASFERITA AD ALTRO ENTE
<b>B</b>	SCELTA DEL MEDICO USL IN REGIONE
<b>C</b>	SCELTA DEL MEDICO FUORI REGIONE
<b>D</b>	ASS.SCAD.DOMICIL.CON RESID.NS.AUSL
<b>E</b>	ELIMINATO A CAUSA DI ERRORE
<b>F</b>	ASSISTENZA TRASF.COMUNICAZ.REGIONE
<b>H</b>	EMIGRAZIONE ESTERO
<b>I</b>	IRREPERIBILITA'
<b>L</b>	LICENZIAMENTO
<b>M</b>	MORTE
<b>O</b>	REVOCA/RIFIUTO TITOLO DI SOGGIORNO
<b>P</b>	ASSIS.SCAD.DOMIC.RESIDENZA NO AUSL
<b>Q</b>	REVOCA PEDIATRA A SEDICENNE
<b>R</b>	EMIGRAZIONE FUORI REGIONE
<b>S</b>	SCELTA MEDICO FUORI-VECCHIE POSIZ.
<b>T</b>	TRASLOCO GENERICO - VECCHIE POSIZ.
<b>U</b>	SCELTA MEDICO IN ALTRO DISTRET.AUSL
<b>V</b>	NON IMPATT. CON ANAG.COMUNE DI RES.
<b>W</b>	ASSIST.SCAD.A STRANIERO RESIDENTE
<b>Y</b>	AS.SCAD.STRAN.EFFET.DIMORA NS AUSL
<b>Z</b>	EMIGRAZIONE IN REGIONE
<b>99</b>	CESSATA ETA' PEDIATRICA

**Tab. 2.2:** Distribuzione record per codice di sospensione dell'assistenza – Anagrafe Sanitaria

<b>COD</b>	<b>2009</b>	
	<b>freq</b>	<b>%</b>
<b>null</b>	1287018	80,1
<b>1</b>	350	0,0
<b>3</b>	107	0,0
<b>4</b>	365	0,0
<b>5</b>	1	0,0
<b>6</b>	2184	0,1
<b>7</b>	11	0,0
<b>8</b>	3	0,0
<b>9</b>	62	0,0
<b>A</b>	341	0,0
<b>B</b>	4141	0,3

<b>C</b>	4259	0,3
<b>D</b>	2316	0,1
<b>E</b>	2121	0,1
<b>F</b>	220	0,0
<b>H</b>	3076	0,2
<b>I</b>	7136	0,4
<b>L</b>	3	0,0
<b>M</b>	145016	9,0
<b>N</b>	2	0,0
<b>O</b>	37	0,0
<b>P</b>	33704	2,1
<b>Q</b>	12	0,0
<b>R</b>	38969	2,4
<b>S</b>	3713	0,2
<b>T</b>	22723	1,4
<b>U</b>	111	0,0
<b>V</b>	742	0,0
<b>W</b>	6772	0,4
<b>Y</b>	5248	0,3
<b>Z</b>	36689	2,3

## 2.2 - Distribuzioni di qualità per gli altri archivi

Per valutare eventuali evidenti discordanze numeriche è stato annotato il totale dei record per anno per i singoli archivi.

**Tab. 2.3:** Distribuzione record per archivio, per anno di disponibilità

ANNO	Mortalità	SDO	Farmaceutica	Esenzione Ticket
1996	6085			13419*
1997	6380			990
1998	6574			2531
1999	6868			2050
2000	6699			2112
2001	6576			1701
2002	6626	138689		1776
2003	6894	133209	5000508	1701
2004	6690	132648	5220546	1697
2005	6759	121444	5246304	1914
2006	6621	121580	5437893	2115
2007	6858	123608	5663056	1850
2008		114597	5989211	1735

\* Anno inizio esenzione minore di 1997

Per i singoli archivi sono state fatte anche valutazioni di qualità più specifiche.

### Mortalità

Poiché per l'applicazione degli algoritmi di stima sono necessarie più fonti relative ad uno stesso periodo, l'archivio di mortalità è stato considerato a partire dall'anno 1996; a questi dati di mortalità si farà riferimento nella trattazione che segue.

**Tab. 2.4:** Distribuzione record per anno e mese - mortalità

	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
<b>Gen</b>	601	712	583	688	651	600	584	607	637	663	622	644
<b>Feb</b>	496	513	562	730	616	535	549	538	643	644	550	589
<b>Mar</b>	511	565	614	616	553	530	618	647	619	615	591	584
<b>Apr</b>	487	512	562	584	557	510	600	539	536	538	529	531
<b>Mag</b>	539	517	523	540	543	546	560	531	557	530	512	547
<b>Giu</b>	442	473	513	521	506	534	581	622	503	560	553	513
<b>Lug</b>	473	471	569	536	541	531	498	520	540	534	588	598
<b>Ago</b>	455	466	505	524	582	562	458	719	485	498	529	539
<b>Set</b>	453	473	488	447	459	529	467	468	481	501	468	527
<b>Ott</b>	539	557	531	573	531	540	579	605	548	552	569	601
<b>Nov</b>	526	543	535	499	574	526	566	511	558	540	572	565
<b>Dic</b>	563	578	589	610	586	633	566	587	583	584	538	620

Si osserva una distribuzione nel complesso uniforme dei record mensili, come atteso.

E' stata inoltre valutata la distribuzione dei record di mortalità per grandi gruppi di cause di morte, per anno (vedasi **Allegato B.1**).

**Tab. 2.5:** Distribuzione record per gruppo di causa di morte e anno - mortalità

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	TOT
1996	30	1930	175	31	170	150	2568	326	240	74	0		27	17	14	52	281	6085
1997	26	1833	172	14	187	170	2710	446	274	113	1	5	25	22	22	66	294	6380
1998	34	1952	130	26	236	178	2765	461	268	91	1	4	18	24	12	49	325	6574
1999	44	2024	181	18	216	180	2844	495	280	103	1	4	29	37	14	65	333	6868
2000	46	1983	190	18	238	188	2820	431	271	92	0	11	28	19	21	34	309	6699
2001	41	2015	186	37	227	207	2742	389	257	96	1	12	18	12	17	51	268	6576
2002	49	2062	145	32	209	179	2793	394	270	103	1	5	25	14	17	32	296	6626
2003	46	1982	214	27	282	225	2847	482	289	91	0	5	30	9	16	33	316	6894
2004	61	2041	195	18	244	178	2738	480	266	95	0	6	29	13	14	45	267	6690
2005	66	1976	213	32	262	191	2692	512	269	122	0	14	34	14	22	55	285	6759
2006	63	2070	184	37	296	198	2544	410	283	126	1	6	37	20	21	76	249	6621
2007	73	2108	227	35	312	222	2634	506	285	96	0	13	27	8	10	76	226	6858
TOT	579	23976	2212	325	2879	2266	32697	5332	3252	1202	6	85	327	209	200	634	3449	79630

Il numero di morti per anno per le diverse cause è piuttosto uniforme, come atteso.

### Schede di dimissione ospedaliera (SDO)

L'archivio SDO comprende la mobilità passiva.

E' stata valutata la distribuzione per anno dei record per grandi gruppi di malattie secondo la classificazione ICD - IX (vedasi **Allegato B.1**).

**Tab. 2.6:** Distribuzione record per gruppo di causa e anno - SDO

	2002	2003	2004	2005	2006	2007	2008	Tot
<b>I</b>	2164	2152	1996	1910	1962	2196	2146	14526
<b>II</b>	13738	13527	14195	13325	12855	13571	11822	93033
<b>III</b>	2996	2832	2811	2548	2384	2655	2288	18514
<b>IV</b>	1191	1128	1290	1189	1309	1381	1502	8990
<b>V</b>	3531	3601	3477	3338	2976	3368	3177	23468
<b>VI</b>	12755	8958	6924	6459	5974	6317	5541	52928
<b>VII</b>	20464	19901	20938	20135	18343	19558	17098	136437
<b>VIII</b>	8781	8640	9081	8776	8009	9110	8273	60670
<b>IX</b>	11976	11225	11744	10802	10504	10749	9689	76689
<b>X</b>	8689	9021	9386	8572	8526	8853	7951	60998
<b>XI</b>	10559	10745	11414	10116	9914	10569	10009	73326
<b>XII</b>	1293	1204	1225	1153	1147	1165	1148	8335
<b>XIII</b>	9040	8918	9825	9445	8859	9414	8138	63639
<b>XIV</b>	1104	1058	1174	1110	1134	1168	1000	7748
<b>XV</b>	1140	1377	1318	1332	1282	1308	1271	9028
<b>XVI</b>	4606	4297	4214	3308	3305	3531	3094	26355
<b>XVII</b>	9636	9217	9213	8564	8187	8533	7513	60863
<b>Codici V</b>	15024	15407	16260	13890	13236	13969	12937	100723
<b>null*</b>	1	0	1	4	0	0	0	6
<b>Tot</b>	138688	133208	136486	125976	119906	127415	114597	896276

\* fanno parte della mobilità passiva extra regionale

## Prescrizioni farmaceutiche

**Tab. 2.7:** Distribuzione record per anno e mese – Prescrizioni farmaceutiche

	2003	2004	2005	2006	2007	2008
<b>Gen</b>	457.536	459.154	447.074	470.718	515.551	565.040
<b>Feb</b>	410.542	419.775	449.244	449.672	465.432	506.527
<b>Mar</b>	429.471	480.675	473.011	500.514	496.392	490.948
<b>Apr</b>	416.336	442.034	418.187	422.521	460.133	510.224
<b>Mag</b>	444.517	452.337	473.704	516.135	524.402	522.677
<b>Giu</b>	398.753	439.474	439.294	461.587	458.949	485.910
<b>Lug</b>	417.128	416.368	398.270	420.404	451.893	500.024
<b>Ago</b>	318.720	348.102	378.161	388.370	395.941	382.997
<b>Set</b>	419.223	418.549	427.083	425.461	422.984	491.392
<b>Ott</b>	449.738	422.282	436.956	466.604	508.329	528.590
<b>Nov</b>	400.712	441.956	457.664	480.108	508.207	486.711
<b>Dic</b>	437.832	479.623	447.656	435.769	454.843	517.897
<b>Tot</b>	5.000.508	5.220.329	5.246.304	5.437.863	5.663.056	5.988.937

Nel complesso non si evidenziano aspetti problematici.

E' stata valutata la distribuzione per anno dei record per classe ATC di prescrizione farmaceutica (vedasi **Allegato B.2**):

**Tab. 2.8:** Distribuzione record per classe ATC e anno – Prescrizioni farmaceutiche

	2003	2004	2005	2006	2007	2008	tot
<b>A</b>	571.750	617.419	642.978	716.070	787.726	891.936	4.227.879
<b>B</b>	341.802	364.318	380.979	405.569	427.153	459.560	2.379.381
<b>C</b>	1.957.881	2.056.568	2.026.019	2.100.381	2.175.452	2.255.446	12.571.747
<b>D</b>	19.423	21.640	21.240	22.187	22.049	23.180	129.719
<b>G</b>	270.292	262.666	243.135	227.801	218.975	220.741	1.443.610
<b>H</b>	179.800	190.560	197.547	208.362	213.290	227.369	1.216.928
<b>J</b>	622.591	618.803	636.550	620.199	664.617	673.637	3.836.397
<b>L</b>	25.552	17.452	13.589	11.750	11.068	11.986	91.397
<b>M</b>	268.900	287.079	269.204	282.390	295.710	323.625	1.726.908
<b>N</b>	396.086	394.220	417.645	445.999	437.621	487.279	2.578.850
<b>P</b>	9.258	9.645	9.849	9.885	10.194	10.673	59.504
<b>R</b>	246.226	293.181	314.258	318.466	338.666	343.226	1.854.023
<b>S</b>	86.729	82.380	69.317	65.052	60.312	60.062	423.852
<b>V</b>	4.218	4.398	3.994	3.752	223	217	16.802
<b>tot</b>	5.000.508	5.220.329	5.246.304	5.437.863	5.663.056	5.988.937	32.556.997

La distribuzione mostra complessivamente una tendenza all'aumento delle prescrizioni.

## Esenzioni Ticket (ET)

**Tab. 2.9:** Distribuzione record per anno di inizio esenzione per totale e per il diabete – Esenzioni Ticket

<b>Anno Inizio Esenzione</b>	<b>Numero Record</b>	<b>Diabete ricodificato</b>
<b>&lt;1997</b>	212.129	13.419
<b>1997</b>	13.137	990
<b>1998</b>	14.448	2.531
<b>1999</b>	56.803	2.050
<b>2000</b>	67.826	2.112
<b>2001</b>	38.849	1.701
<b>2002</b>	17.462	1.776
<b>2003</b>	16.376	1.701
<b>2004</b>	16.310	1.697
<b>2005</b>	16.283	1.914
<b>2006</b>	15.395	2.115
<b>2007</b>	65.833	1.850
<b>2008</b>	23.692	1.735

L'archivio dell'esenzioni ticket è un database unico (non suddiviso per anni): contiene tutte le esenzioni rilasciate dall'Ausl fino ad oggi.

I codici delle esenzioni hanno subito delle modifiche negli anni, si nota, infatti, che nel 1999/2000 e nel 2007 il n° di esenzioni sembrano aumentare improvvisamente, in realtà un soggetto esente è presente nella banca dati più volte con diversi codici di esenzione per una stessa patologia. Per avere delle statistiche di incidenza/prevalenza corrette si è provveduto a ricodificare le patologie. Per esempio il diabete, che negli anni ha avuto i seguenti codici: 013, 013.250, DIA e DIM, ricodificando i record con un unico codice la distribuzione negli anni si è uniformata.

Inoltre, la data di inizio esenzione può corrispondere alla prima esenzione per la stessa causa o alla data del rinnovo dell'esenzione. ed, in alcuni casi, è chiaramente fittizia e posta in epoca prossima a inizio 1900 (probabilmente in epoca istituzione archivio informatizzato mancava l'informazione data inizio esenzione per quelle già in essere).

### **Fase 3: record linkage degli archivi con l'Anagrafe Sanitaria**

Tempo di realizzazione: 1 Luglio – 30 Settembre 2009

#### *3.1 - Valutazione storicizzazione dell'archivio anagrafico*

L'archivio anagrafico non è storicizzato ma si è potuto effettuare un'analisi della completezza dell'anagrafica attraverso l'incrocio con fotografie della banca dati del 2008, 2007, 2006, 2005. L'anagrafe attuale (al 2009) contiene il 100% dei record degli anni precedenti; sono possibili, invece, nuovi assistiti nel 2009 non presenti nelle fotografie degli anni precedenti (nuovi nati, immigrati).

#### *3.2 - Individuazione e valutazione delle chiavi di linkage utilizzabili per singolo archivio*

Tutti gli archivi a nostra disposizione, tranne la mortalità, ci sono stati forniti con il campo chiave RNLIB che ci permette il linkage con l'anagrafe e fra archivi sanitari.

Il campo chiave dell'archivio anagrafico RNLIB viene individuato dal programma della medicina di base che lo rende disponibile a tutti gli altri flussi aziendali. Per la parte dei dati che arriva dall'esterno (es. mobilità passiva, farmaceutica) si effettua un record linkage tramite codice fiscale, cognome, nome, data di nascita; per i record non linkati si provvede ad un controllo manuale uno a uno.

Solo per l'archivio della mortalità si è dovuto procedere al record linkage con l'anagrafe sanitaria per recuperare la chiave RNLIB. Sono state individuate le variabili da utilizzare nel linkage, per impostare l'algoritmo specifico; si è quindi valutata la completezza dei campi stessi.

Le chiavi di linkage individuate sono le seguenti:

- ❖ **Tessera sanitaria (TS)**
- ❖ **Codice fiscale (CF)**
- ❖ **Nome**
- ❖ **Cognome**
- ❖ **Data di nascita**

Sono stati valutati per ogni fonte la completezza, per anno e nel totale, delle singole chiavi di linkage ed il loro formato.

## Mortalità

**Tab. 3.1:** Percentuale di completezza chiavi di linkage per anno - mortalità

chiavi	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
<b>TS</b>	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
<b>CF</b>												
formato testo	0%	0%	0%	78%	78%	78%	100%	100%	100%	100%	100%	100%
<b>nome</b>												
formato testo	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
<b>cognome</b>												
formato testo	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%
<b>data nascita</b>												
formato data/ora	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

## Schede di dimissione ospedaliera (SDO)

**Tab. 3.2:** Percentuale di completezza chiave di linkage per anno – SDO

SDO	RNLIB mancante		
	n°	%	% non residenti
<b>2002</b>	7644	5,5	81,9
<b>2003</b>	7702	5,8	81,7
<b>2004</b>	7620	5,6	77,9
<b>2005</b>	1155	0,9	0,0
<b>2006</b>	911	0,8	0,0
<b>2007</b>	917	0,7	0,0
<b>2008</b>	406	0,4	0,0
<b>Totale</b>	26355	2,9	

## Prescrizioni farmaceutiche

**Tab. 3.3:** Percentuale di completezza chiave di linkage per anno – Prescrizioni farmaceutiche

FT	RNLIB mancante		
	n°	%	% non residenti
2003	121564	2,4	16,4
2004	112191	2,1	47,1
2005	157895	3,0	36,9
2006	153172	2,8	27,0
2007	112991	2,0	-
2008	86315	1,4	-
<b>Totale</b>	744128	2,3	

La procedura utilizzata per il record linkage dell'archivio mortalità è di tipo semi-deterministico, ovvero un record è considerato riconducibile al medesimo soggetto solo se la corrispondenza fra campi riferiti alla stessa variabile è perfetta, con una procedura a cascata di passi sempre meno restrittivi.

Le informazioni anagrafiche contenute in tutti gli archivi sono state utilizzate per identificare i soggetti nell'Anagrafe mediante i seguenti 5 passi di chiave:

- step 1 : codice fiscale, cognome, nome, data di nascita
- step 2 : codice fiscale, cognome, data di nascita
- step 3 : codice fiscale, cognome, nome
- step 4 : codice fiscale, nome, data di nascita
- step 5 : cognome, nome, data di nascita

### *3.3 - Esecuzione di procedure di record linkage semi-deterministico a cascata*

E' stata elaborata una procedura per effettuare il record linkage in Microsoft SQL Server 2005, con l'obiettivo di massimizzare standardizzazione (replicabilità) e sicurezza.

La procedura è stata elaborata sostanzialmente de novo.

Sono state formulate una serie di query eseguite a cascata; la logica di ogni ciclo di query è selezionare i soggetti che linkano per il passo di chiave, sottrarli dall'archivio con il quale si è effettuato il linkage e creare il nuovo archivio dopo la sottrazione, oltre ad una tabella contenente i record linkati. Al termine dello step "n" si otterranno dunque due tabelle:

- tab\_stepn: contiene i record linkati nello step n
- tab\_dopo\_stepn: contiene i record dell'archivio che non sono ancora stati linkati fino a questo punto (è la nuova tabella che verrà incrociata con l'anagrafe allo step n+1)

Nella query di incrocio viene aggiunta ad ogni record la variabile RNLIB, ottenuta dall'Anagrafe.

Dall'unione di tutte le tabelle tab\_stepn si ottiene il dataset con tutti i record agganciati; in questo dataset si eliminano le variabili anagrafiche dell'archivio, poiché le informazioni anagrafiche utilizzate in seguito saranno esclusivamente derivate dall'Anagrafe Sanitaria.

### 3.4 - Valutazione delle percentuali di linkage

Si riporta di seguito la tabella con il numero assoluto di record linkati per singolo passo di chiave, il numero degli appaiati cumulati nonché dei record non agganciati, con le relative percentuali. La quota di non appaiati è molto bassa pari a 2,6% per tutto il periodo

#### Mortalità

**Tab. 3.4:** Percentuali di incrocio per singolo passo di chiave e totale - mortalità

<i>Fonte</i>	<i>Periodo</i>	<i>Passo di chiave</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
mortalità	1996-2007	K0	<b>88001</b>			
		CFCNDN	59755	59755	67,9%	67,9%
		CFCDN	214	59969	0,2%	68,1%
		CFCN	119	60088	0,1%	68,3%
		CFNDN	241	60329	0,3%	68,6%
		CNDN	25420	85749	28,9%	97,4%

**Legenda:**

**CF:** Tessera Sanitaria

**C:** Cognome

**N:** Nome

**DN:** Data di Nascita

## Fase 4: produzione di stime di frequenza

Tempo di realizzazione: settembre – novembre 2009

### 4.1 - Applicazione agli archivi di algoritmi forniti

Gli algoritmi applicati alla popolazione dell'AUSL Modena corrispondono a quelli presentati nel Supplemento della Rivista "Epidemiologia e Prevenzione", Anno 32 (3), maggio-giugno 2008. Per le stime della Cardiopatia ischemica non si è potute utilizzare l'archivio delle esenzioni ticket in quanto il codice di esenzione per questa patologia (002.414) non viene utilizzato.

### Risultati

IMA

**Tab.4.1-IMA:** Incidenza grezza e standardizzata dell'IMA (\*100 000 abitanti)

anno	sezzo	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2007	F	579	341236	<b>169,68</b>	<b>121,96</b>
2007	M	879	328862	<b>267,29</b>	<b>274,30</b>
2007	T	1458	670098	<b>217,58</b>	<b>191,39</b>

**Tab. 4.2-IMA:** Distribuzione casi per classi d'età e anno

Classe età	anno
<b>Frequenza Percent col</b>	<b>2007</b>
15-34	3 0,21
35-54	174 11,93
55-64	217 14,88
65-74	323 22,15
75+	741 50,82
<b>Totale</b>	<b>1458</b>

**Graf. 4.2-IMA:** Contributo esclusivo delle fonti per anno di stima

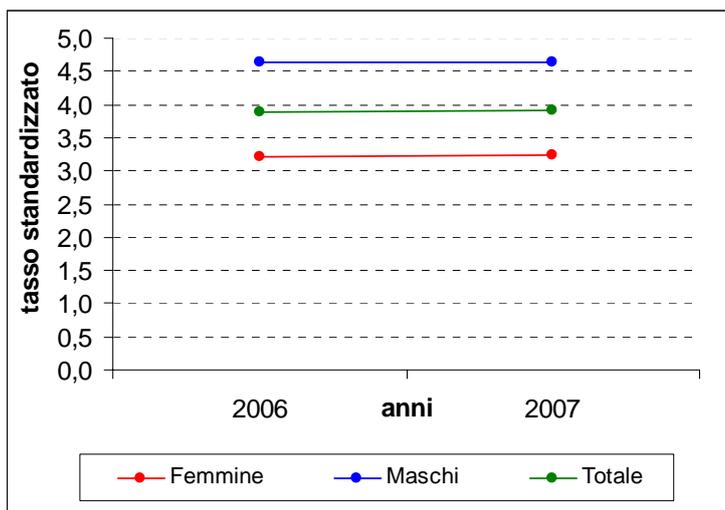
Contributo esclusivo fonti		
Fonte	2007	
	Freq.	%
<b>SDO</b>	1242	85,19
<b>CM</b>	216	14,81
<b>totale</b>	1458	100

## DIABETE

**Tab. 4.1-diabete:** Prevalenza grezza e standardizzata del diabete (\*100 abitanti)

anno	sezzo	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2006	F	12815	338716	3,78	<b>3,21</b>
2006	M	14900	326651	4,56	<b>4,64</b>
2006	totale	27715	665367	4,17	<b>3,89</b>
2007	F	13071	341236	3,83	<b>3,24</b>
2007	M	15042	328862	4,57	<b>4,63</b>
2007	totale	28113	670098	4,20	<b>3,90</b>

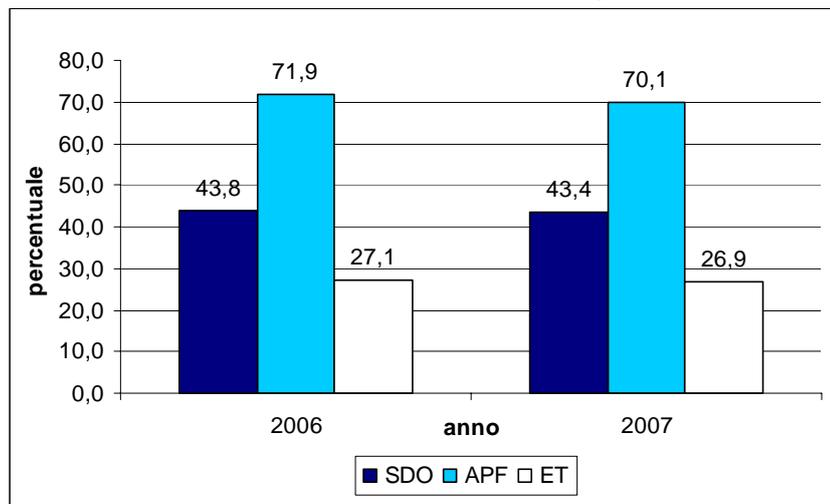
**Graf. 4.1-diabete:** Prevalenza standardizzata del diabete (\*100 abitanti)



**Tab. 4.2-diabete:** Distribuzione casi per classi d'età e anno di stima

classe età	anno		
	2006	2007	Totale
<b>0-14</b>	85 0,3	96 0,3	181
<b>15-34</b>	505 1,8	509 1,8	1014
<b>35-54</b>	4148 15,0	4165 14,8	8313
<b>55-64</b>	6126 22,1	6179 22,0	12305
<b>65-74</b>	8239 29,7	8312 29,6	16551
<b>75+</b>	8612 31,1	8852 31,5	17464
<b>Totale</b>	27715	28113	55828

**Graf. 4.2-diabete:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3-diabete:** Contributo esclusivo delle fonti, per anno di stima

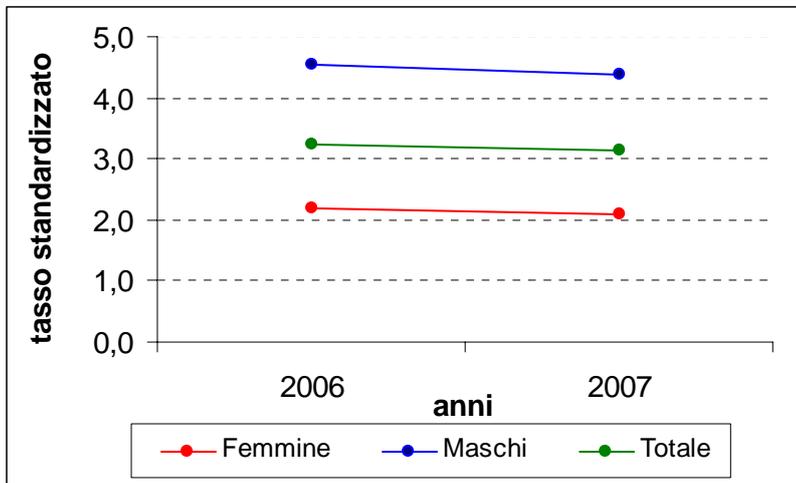
Contributo esclusivo fonti				
Fonte	2006		2007	
	Freq.	%	Freq.	%
<b>SDO</b>	4.410	15,9	4.777	17,0
<b>APF</b>	10.020	36,2	10.249	36,5
<b>ET</b>	2.619	9,4	2.816	10,0
<b>più di una fonte</b>	10.666	38,5	10.271	36,5
<b>Totale</b>	27.715	100	28.113	100

## CARDIOPATIA ISCHEMICA

**Tab. 4.1-cardiopatía ischemica:** Prevalenza grezza e standardizzata della cardiopatía ischemica (\*100 abitanti)

anno	sesto	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2006	F	10055	338716	2,97	<b>2,19</b>
2006	M	14159	326651	4,33	<b>4,53</b>
2006	totale	24214	665367	3,64	<b>3,24</b>
2007	F	9853	341236	2,89	<b>2,09</b>
2007	M	13995	328862	4,26	<b>4,39</b>
2007	totale	23848	670098	3,56	<b>3,14</b>

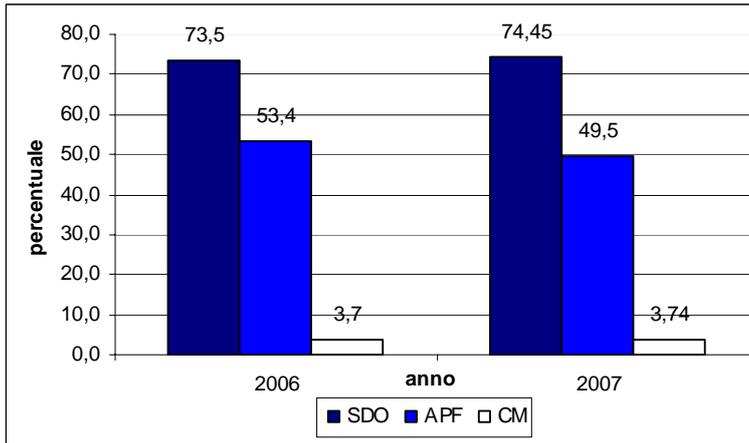
**Graf. 4.1-cardiopatía ischemica:** Prevalenza standardizzata della cardiopatía ischemica (\*100 abitanti)



**Tab. 4.2-cardiopatía ischemica:** Distribuzione casi per classi d'età e anno di stima

Classe età	anno		Totale
	2006	2007	
<b>0-14</b>	7 0,03	8 0,03	15
<b>15-34</b>	71 0,29	61 0,26	132
<b>35-54</b>	1515 6,26	1509 6,33	3024
<b>55-64</b>	3310 13,67	3276 13,74	6586
<b>65-74</b>	6208 25,64	6064 25,43	12272
<b>75+</b>	13103 54,11	12930 54,22	26033
<b>Totale</b>	24214	23848	48062

**Graf. 4.2-cardiopatia ischemica:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- cardiopatia ischemica:** Contributo esclusivo delle fonti per anno di stima

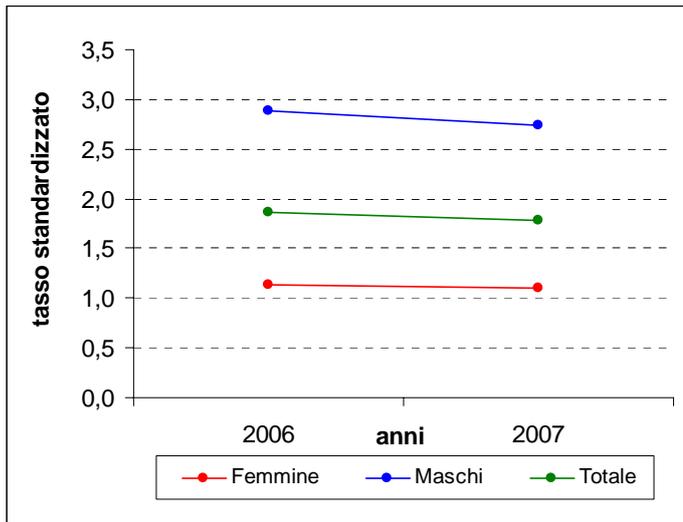
Contributo esclusivo fonti				
Fonte	2006		2007	
	Freq.	%	Freq.	%
<b>SDO</b>	10660	44,0	11415	47,9
<b>APF</b>	6130	25,3	5741	24,1
<b>CM</b>	246	1,0	290	1,2
<b>più di una fonte</b>	7178	29,6	6402	26,8
<b>totale</b>	24214	100,0	23848	100,0

## BPCO

**Tab. 4.1-BPCO:** Prevalenza grezza e standardizzata della BPCO, 35+ anni (\*100 abitanti)

anno	secco	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2006	F	3000	220171	1,4	1,1
2006	M	5276	200118	2,6	2,9
2006	totale	8276	420289	2,0	1,9
2007	F	2963	223149	1,3	1,1
2007	M	5095	203197	2,5	2,7
2007	totale	8058	426346	1,9	1,8

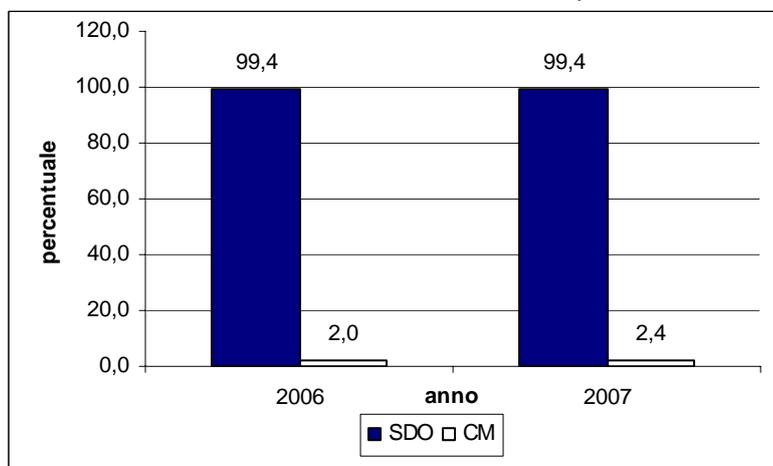
**Graf. 4.1-BPCO:** Prevalenza standardizzata della BPCO, 35+ anni (\*100 abitanti)



**Tab. 4.2-BPCO:** Distribuzione casi per classi d'età e anno di stima

Classe età	anno		Totale
	2006	2007	
35-54	494 5,97	454 5,63	948
55-64	915 11,06	909 11,28	1824
65-74	2180 26,34	2089 25,92	4269
75+	4687 56,63	4606 57,16	9293
<b>Totale</b>	<b>8276</b>	<b>8058</b>	<b>16334</b>

**Graf. 4.2-BPCO:** Contributo assoluto delle fonti per anno di stima



Tab. 4.3- BPCO: Contributo esclusivo delle fonti per anno di stima

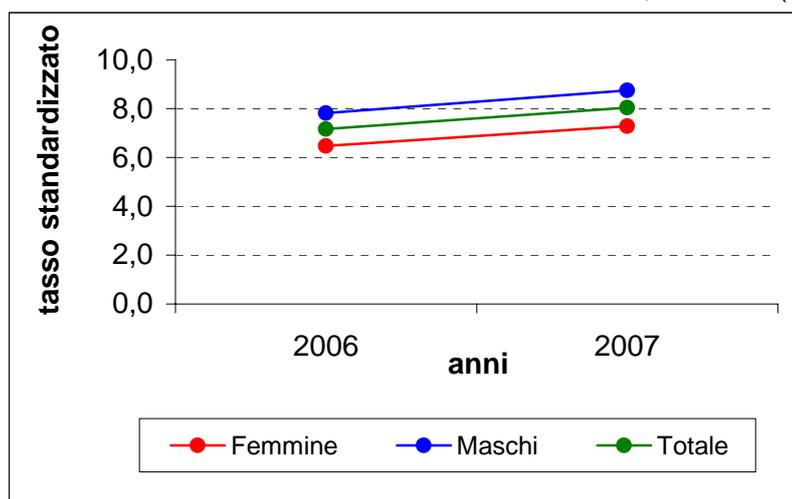
Contributo esclusivo fonti - 35 e più				
Fonte	2006		2007	
	Freq.	%	Freq.	%
<b>SDO</b>	8109	98,0	7867	97,6
<b>CM</b>	46	0,6	47	0,6
<b>più di una fonte</b>	121	1,5	144	1,8
<b>totale</b>	8276	100,0	8058	100,0

ASMA

Tab. 4.1-asma: Prevalenza grezza e standardizzata dell'asma, 0-34 anni (\*100 abitanti)

anno	sexso	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2006	F	8063	118545	6,80	<b>6,47</b>
2006	M	10310	126533	8,15	<b>7,82</b>
2006	totale	18373	245078	7,50	<b>7,17</b>
2007	F	9137	118087	7,74	<b>7,28</b>
2007	M	11657	125665	9,28	<b>8,74</b>
2007	totale	20794	243752	8,53	<b>8,03</b>

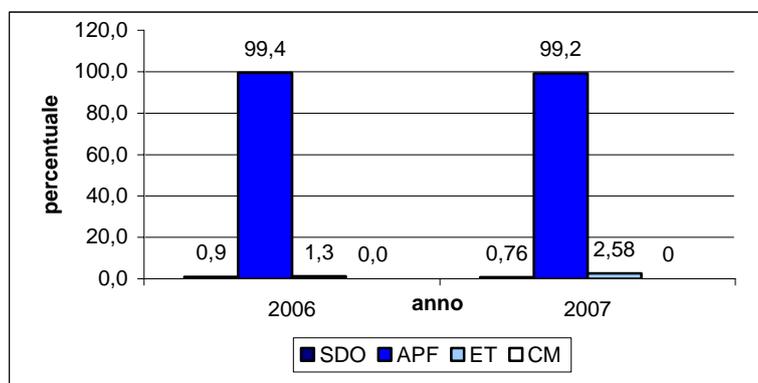
Graf. 4.1-asma: Prevalenza standardizzata dell'asma, 0-34 anni (\*100 abitanti)



**Tab. 4.2-asma:** Distribuzione casi per classi d'età e anno di stima

Classe età	anno		Totale
	2006	2007	
<b>0-14</b>	12569 68,41	14157 68,08	26726
<b>15-34</b>	5804 31,59	6637 31,92	12441
<b>Totale</b>	18373	20794	39167

**Graf. 4.2-asma:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- asma:** Contributo esclusivo delle fonti, per anno di stima

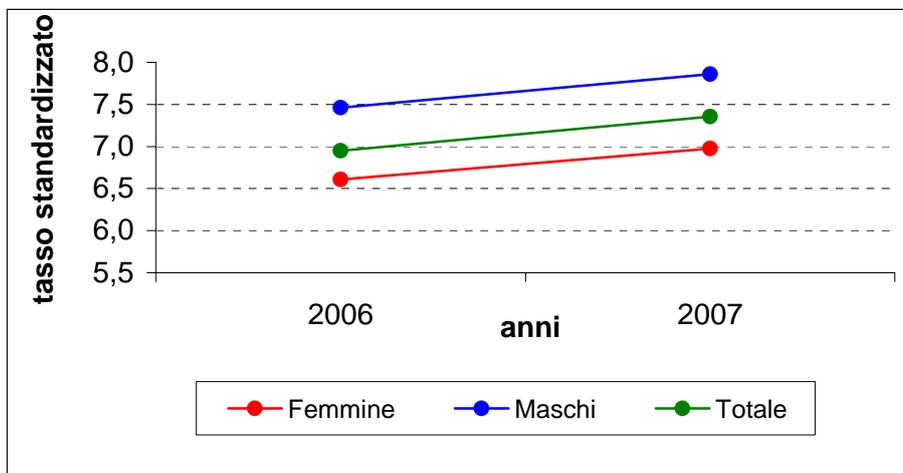
Fonte	2006		2007	
	Freq.	%	Freq.	%
<b>SDO</b>	35	0,2	45	0,22
<b>APF</b>	17.979	97,9	20.116	96,7
<b>ET</b>	67	0,4	122	0,59
<b>CM</b>	-	0,0	-	0
<b>più di una fonte</b>	292	1,6	511	2,46
<b>totale</b>	18.373	100	20.794	100

MPCO

**Tab. 4.1-MPCO:** Prevalenza grezza e standardizzata delle MPCO (\*100 abitanti)

anno	secco	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand
2006	F	23385	338716	6,90	<b>6,61</b>
2006	M	24036	326651	7,36	<b>7,46</b>
2006	totale	47421	665367	7,13	<b>6,95</b>
2007	F	24699	341236	7,24	<b>6,97</b>
2007	M	25860	328862	7,86	<b>7,86</b>
2007	totale	50559	670098	7,55	<b>7,35</b>

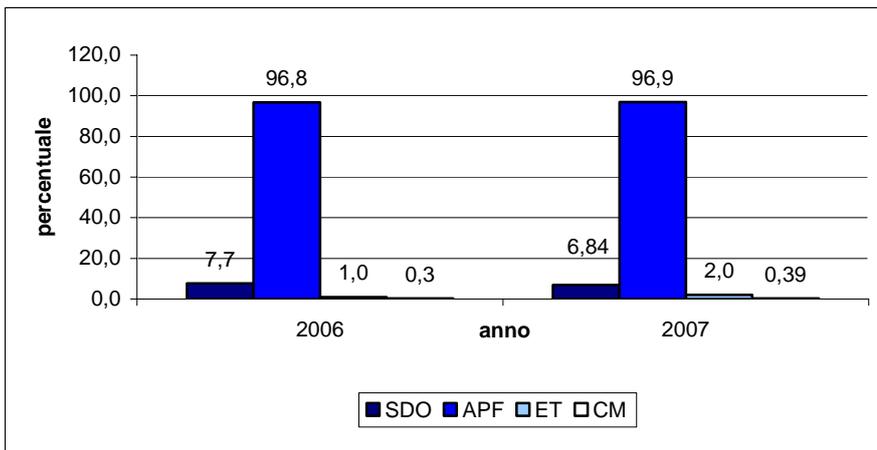
**Graf. 4.1-MPCO:** Prevalenza standardizzata delle MPCO (\*100 abitanti)



**Tab. 4.2-MPCO:** Distribuzione casi per classi d'età e anno di stima

Classe età	anno		Totale
	2006	2007	
<b>0-14</b>	12681 26,74	14255 28,19	26936
<b>15-34</b>	5992 12,64	6833 13,51	12825
<b>35-54</b>	8820 18,6	9662 19,11	18482
<b>55-64</b>	4524 9,54	4821 9,54	9345
<b>65-74</b>	6198 13,07	6408 12,67	12606
<b>75+</b>	9206 19,41	8580 16,97	17786
<b>Totale</b>	47421	50559	97980

**Graf. 4.2-MPCO:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- MPCO:** Contributo esclusivo delle fonti, anni di stima 2004-2008

Fonte	2006		2007	
	Freq.	%	Freq.	%
<b>SDO</b>	1.310	2,8	1.273	2,52
<b>APF</b>	43.248	91,2	46.020	91,0
<b>ET</b>	111	0,2	209	0,41
<b>CM</b>	42	0,1	54	0,1
<b>più di una fonte</b>	2.710	5,7	3.003	5,94
<b>totale</b>	47.421	100	50.559	100

#### 4.2 - Confronto con stime ottenute applicando gli stessi algoritmi in altre realtà italiane

Le altre realtà italiane prese come riferimento per il confronto delle stime di incidenza e prevalenza delle patologie di cui sopra sono quelle presenti nel Supplemento della Rivista "Epidemiologia e Prevenzione" sopra citato. Le stime pubblicate in questo documento sono state prodotte per gli anni 2006 e 2007. Le seguenti considerazioni si riferiscono a questo periodo temporale.

Per quanto riguarda l'**IMA**, le stime di incidenza ottenute a Modena sono più elevate rispetto alle altre realtà italiane, in entrambi i sessi.

La prevalenza stimata di **diabete** a Modena totale e per entrambi i sessi risulta in linea con le altre realtà italiane. Il contributo assoluto della fonte di Modena è molto simile rispetto alle realtà di Torino e Firenze. Il contributo esclusivo della fonte SDO è in generale più alta

Anche se la fonte ET non si è potuta impiegare, perché a Modena il codice di esenzione nel calcolo delle stime della **cardiopatia ischemica** non viene utilizzato, comunque i tassi stimati a Modena per la cardiopatia ischemica sono in generale più in linea con Firenze e Venezia.

A Modena i tassi di prevalenza di **BPCO** negli uomini sono più elevati solo rispetto a Venezia, mentre per le donne sono più bassi di tutti gli altri centri. La distribuzione dei contributi, sia assoluti che esclusivi, è comparabile a quella degli altri centri.

La prevalenza stimata di **asma** a Modena è notevolmente più elevata rispetto a tutti i centri.

Anche le stime di prevalenza di **MPCO** sono più alte a Modena. La distribuzione dei contributi è in linea con gli altri centri.

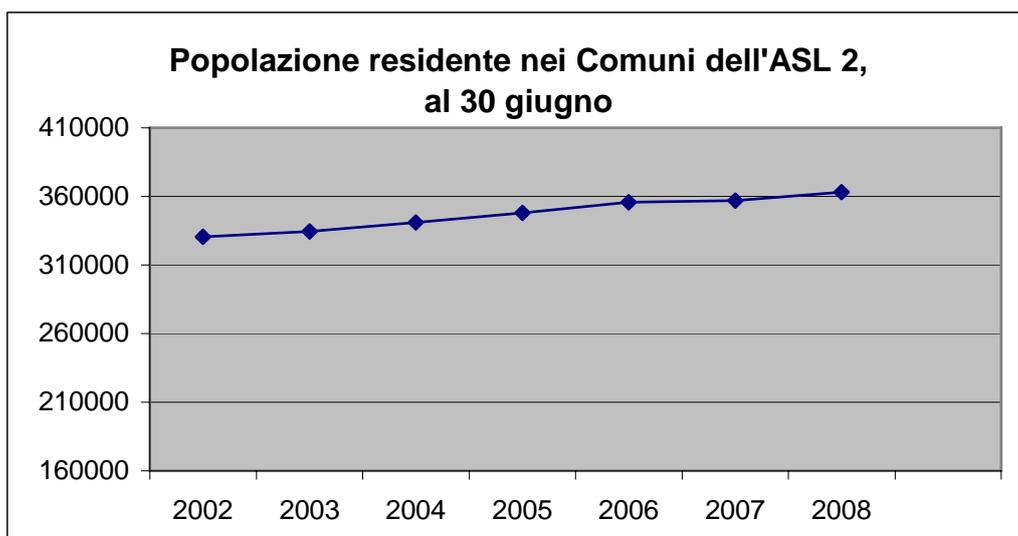
## ASL 2 Perugia

Il territorio dell'Azienda Sanitaria Locale 2 di Perugia comprende 24 Comuni, tra i quali Perugia, capoluogo di provincia.

La seguente tabella riporta la popolazione residente in questi 24 Comuni al 30 giugno degli anni dal 2002 al 2008, per i quali è stato possibile produrre stime di frequenza:

anno	totale
2002	330564
2003	334319
2004	340976
2005	347870
2006	355815
2007	357008
2008	362915

Fonte: ISTAT ([www.demo.istat.it](http://www.demo.istat.it))



Si evidenzia una tendenza all'aumento della popolazione.

## **Fase 1: sistemazione archivi**

Tempo di realizzazione: giugno 2009 - gennaio 2010

### ***1.1 - Acquisizione degli archivi aziendali richiesti per il progetto***

Gli archivi dell'ASL 2 di Perugia disponibili al 1 febbraio 2010 sono i seguenti:

**Tab. 1.1:** Archivi ASL 2 richiesti per il progetto; sono specificati primo e ultimo anno disponibile ed il formato dei dati

<b>FONTE</b>	<b>DA (anno)</b>	<b>A (anno)</b>	<b>FORMATO</b>
ANAGRAFE SANITARIA	2009	2010	Vista su database Oracle a partire dal giugno 2009; file txt su mesi specifici per gli anni 2006, 2007, 2008
ANAGRAFE COMUNALE	archivio non disponibile		
MORTALITA'	1994	2007	Testo, delimitato
SCHEDE DIMISSIONI OSPEDALIERE (SDO)	1998	2008	Vista su database Oracle
MOBILITA' PASSIVA INTRAREGIONE DELLE SDO	1998	2008	Vista su database Oracle
MOBILITA' PASSIVA EXTRAREGIONE DELLE SDO	2000	2008	Vista su database Oracle (dati non disponibili per gli anni 2003 e 2004)
PRESCRIZIONI FARMACEUTICHE	2005	2008	testo, lunghezza fissa
ESENZIONI TICKET	2010	2010	Vista su database Oracle con storicizzazione delle date inizio e fine dell'esenzione
ANATOMIA PATOLOGICA	archivio non disponibile		

Tutti gli archivi sono stati acquisiti ed organizzati in tabelle di Microsoft SQL 2005.

### ***1.2 - Organizzazione dei dati in un unico dataset per archivio***

Gli archivi dell'anagrafe sanitaria e delle esenzioni sono stati ottenuti da una vista contenente la situazione completa a febbraio 2010.

Gli archivi delle schede di dimissione sono stati ottenuti da viste separate per anno.

Tali viste presentavano lo stesso tracciato per cui non si sono riscontrati problemi di popolamento di una tabella unica, contenente i dati di tutti gli anni oggetto di analisi.

L'archivio della mortalità è stato ottenuto da un unico file txt contenente tutti gli anni.

L'archivio delle prescrizioni farmaceutiche è stato costruito a partire da un file txt per il 2005 con tracciato diverso rispetto a quello dei file annuali forniti per gli anni successivi. L'uniformazione dei 2 tracciati non ha però evidenziato particolare problemi.

La sistemazione di ogni archivio ha richiesto un lavoro differente, in quanto sono emerse problematiche diverse.

Alcune problematiche specifiche riscontrate a Perugia:

→ Mortalità: date in un formato testo presenti nel txt da trasformare in un formato data gg/mm/aaaa accettato da MSSQL; mancanza disponibilità dati per l'anno 2008

→ Anagrafe Sanitaria: mancanza degli archivi storicizzati per gli anni precedenti al 2009 per cui il linkage è stato eseguito con l'anagrafe completa a febbraio 2010; sono stati importati file txt di specifici mesi (2006/03-2007/01-2008/04-2009/05) per la fase "record linkage degli archivi con l'Anagrafe Sanitaria". I file txt avevano un tracciato diverso rispetto alla vista disponibile a partire dal 2009/06.

→ Esenzioni: mancanza degli archivi storicizzati per cui le situazioni degli esenti sono state ricostruite a partire dalle date inizio/fine esenzione rilevate nella vista aggiornata a febbraio 2010

→ Mobilità Passiva extra-regione: mancanza disponibilità dati 2003 e 2004; incompletezza dati anagrafici nel 2002

## **Fase 2: valutazione della qualità dei dati**

Tempo di realizzazione: giugno 2009 - gennaio 2010

### **2.1 - Distribuzioni di qualità per l'Anagrafe Sanitaria**

La chiave dell'archivio disponibile a partire dal 2009-06 è la variabile "Codice Assistito Aziendale" che rappresenta l'identificativo aziendale univoco dell'assistibile.

Per quanto riguarda campi importanti per il record linkage e per analisi successive, è stata valutata la completezza annuale come percentuale dei campi compilati sul totale dei record.

La completezza del campo "Codice Fiscale" è attorno all'100% (nell'anagrafe sanitaria 2010-02, su un totale record di 472.807 solo 57 record presentano codice fiscale non valorizzato).

E' stata calcolata la completezza anche dei campi "Cognome", "Nome", "Data di nascita", "comune di residenza", che è attorno al 100% (nell'anagrafe sanitaria 2010-02, un solo record su 472.807 non ha il campo "Cognome" valorizzato; 54 record su 472.807 non hanno il campo "Codice Comune Residenza" valorizzato).

Nel corso dei controlli sono emersi assistiti duplicati, ovvero soggetti con diversa Tessera Sanitaria, ma con stesso CF o più variabili anagrafiche. Inoltre, dato uno stesso CF sono stati rilevati Codici Assistito Aziendale duplicati, legati al cambiamento delle altre variabili anagrafiche.

Si è approfondita la ricerca dei duplicati e sono stati considerati tali i record con corrispondenza completa in una delle seguenti combinazioni di variabili anagrafiche:

6. codice fiscale, cognome, nome, data di nascita
7. codice fiscale, cognome, data di nascita
8. codice fiscale, cognome, nome
9. codice fiscale, nome, data di nascita
10. nome, cognome, data di nascita

I soggetti duplicati (96 record nell'Anagrafe 2010-02) sono stati analizzati manualmente per cercare di identificare i duplicati reali ed eventuali altri problemi associati a questi record.

Si è quindi deciso di mantenere, fra due duplicati, il record con il progressivo di TS più alto.

Si è ritenuto di valutare l'attendibilità del campo "Codice sospensione", per il quale si riportano legenda e distribuzione per gli anni disponibili:

**Tab. 2.1:** Legenda codici di sospensione

<b>Codice</b>	<b>Descrizione</b>
Ø	deroga interna
1	residente
2	non residente in deroga
3	straniero
4	domiciliato
5	residente in deroga
6	trasferito
7	deceduto
8	occasionale
9	errore o altro

**Tab. 2.2:** Distribuzione record per anno e codice di sospensione dell'assistenza – Anagrafe Sanitaria

COD	2006		2007		2008		2009		2010	
	freq	%	freq	%	freq	%	freq	%	freq	%
0	4061	1,00%	4139	1,00%	3552	0,83%	3642	0,84%	3727	0,79%
1	321651	79,02%	322899	78,00%	322789	75,81%	323478	74,59%	324895	68,72%
2	206	0,05%	208	0,05%	209	0,05%	197	0,05%	1174	0,25%
3	36402	8,94%	39052	9,43%	45281	10,63%	47264	10,90%	49775	10,53%
4	4255	1,05%	4593	1,11%	5299	1,24%	6048	1,39%	6578	1,39%
5	1995	0,49%	1908	0,46%	2847	0,67%	2389	0,55%	3805	0,80%
6	9674	2,38%	9610	2,32%	9612	2,26%	10014	2,31%	38408	8,12%
7	28814	7,08%	31555	7,62%	36216	8,51%	40617	9,37%	43445	9,19%
8									151	0,03%
9									849	0,18%
<b>Tot</b>	<b>407058</b>	<b>100,00%</b>	<b>413964</b>	<b>100%</b>	<b>425805</b>	<b>100%</b>	<b>433649</b>	<b>100%</b>	<b>472807</b>	<b>100,00%</b>

## 2.2 - Distribuzioni di qualità per gli altri archivi

Per valutare eventuali evidenti discordanze numeriche è stato annotato il totale dei record per anno per i singoli archivi, senza eseguire il linkage con l'anagrafe sanitaria

**Tab. 2.3:** Distribuzione record per archivio, per anno di disponibilità

ANNO	Mortalità	SDO	Mobilità Passiva		Farmaceutica	Esenzioni Ticket
			Intra	Extra		
1998	3637	25183	51057			
1999	3489	24226	50299			
2000	3402	23132	51646	4876		
2001	3522	21644	49406	5175		
2002	3537	20539	47158	5276		
2003	3708	21205	45384			
2004	3320	20398	45472			
2005	3705	20101	42819	5888	3852011	
2006	3524	19684	43310	5884	4171574	
2007	3537	19503	43035	6215	4394230	
2008		18888	42181	6458	4636784	
2009						
2010						213917
<b>Tot</b>	<b>35381</b>	<b>234503</b>	<b>511767</b>	<b>39772</b>	<b>17054599</b>	<b>213917</b>

Per i singoli archivi sono state fatte anche valutazioni di qualità più specifiche.

### Mortalità

Poiché per l'applicazione degli algoritmi di stima sono necessarie più fonti relative ad uno stesso periodo, l'archivio di mortalità è stato considerato a partire dall'anno 1998; a questi dati di mortalità si farà riferimento nella trattazione che segue.

**Tab. 2.4:** Distribuzione record per anno e mese - mortalità

	<b>1998</b>	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>2004</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>	<b>TOT</b>
<b>gen</b>	297	393	376	341	401	340	297	340	331	315	
<b>feb</b>	328	315	308	281	298	325	300	414	316	278	
<b>mar</b>	370	332	280	323	349	375	304	361	291	292	
<b>apr</b>	320	296	258	292	288	272	269	299	275	312	
<b>mag</b>	306	287	260	304	291	279	265	274	280	261	
<b>giu</b>	274	238	248	261	310	294	264	284	282	280	
<b>luglio</b>	295	256	292	262	245	325	270	284	283	326	
<b>ago</b>	271	254	276	274	220	335	219	224	263	274	
<b>set</b>	259	261	242	260	243	256	242	281	248	274	
<b>ott</b>	287	260	288	295	310	299	293	314	327	329	
<b>nov</b>	319	275	270	280	308	281	278	300	303	286	
<b>dic</b>	311	322	304	349	274	327	319	330	325	310	
<b>tot</b>	<b>3637</b>	<b>3489</b>	<b>3402</b>	<b>3522</b>	<b>3537</b>	<b>3708</b>	<b>3320</b>	<b>3705</b>	<b>3524</b>	<b>3537</b>	<b>35381</b>

Si osserva una distribuzione nel complesso uniforme dei record mensili, come atteso.

E' stata inoltre valutata la distribuzione dei record di mortalità per grandi gruppi di cause di morte, per anno (vedasi **Allegato B.1**).

**Tab. 2.5:** Distribuzione record per gruppo di causa di morte e anno - mortalità

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	tot
<b>1998</b>	3	987	60	12	36	91	1796	254	166	52		5	2	8	4	33	128	3637
<b>1999</b>	3	1005	52	14	34	68	1695	216	135	53		6	4	9	14	40	141	3489
<b>2000</b>	2	967	60	12	40	93	1660	203	116	57		12	5	4	4	31	136	3402
<b>2001</b>	0	931	45	18	31	86	1825	234	143	42		12	3	2	6	38	106	3522
<b>2002</b>	0	1015	34	15	21	99	1701	237	116	59		10	3	6	6	87	128	3537
<b>2003</b>	1	1060	52	9	53	124	1730	266	144	48	1	19	9	8	10	49	125	3708
<b>2004</b>	1	984	42	14	43	74	1523	261	117	60		14	6	9	5	29	138	3320
<b>2005</b>	3	1086	61	16	42	128	1692	313	117	57		19	10	8	3	36	114	3705
<b>2006</b>	1	1057	47	12	80	146	1561	237	120	50		13	19	9	2	51	119	3524
<b>2007</b>	4	1029	81	20	74	180	1526	259	87	41		22	15	4	2	53	140	3537
<b>tot</b>	18	10121	534	142	454	1089	16709	2480	1261	519	1	132	76	67	56	447	1275	35381

Il numero di morti per anno per le diverse cause è piuttosto uniforme, come atteso.

#### Schede di dimissione ospedaliera (SDO)

Anche per questo archivio è stata valutata la distribuzione per anno dei record per grandi gruppi di malattie secondo la classificazione ICD - IX (vedasi **Allegato B.1**), eseguendo il linkage con l'anagrafe sanitaria 2010-02

**Tab. 2.6:** Distribuzione record per gruppo di causa e anno - SDO

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XIII	XIV	XV	XVI	XVII	codici v	tot
<b>1998</b>	377	1759	874	248	534	401	2377	933	2687	1465	1355	354	373	107	433	1384	834	1522	18017
<b>1999</b>	395	1861	809	271	521	493	2361	941	2722	1411	1501	404	328	94	428	1115	691	1339	17685
<b>2000</b>	323	1886	761	232	441	714	2416	894	2523	1533	1511	381	416	132	382	1043	688	1220	17496
<b>2001</b>	314	1783	461	222	442	779	2203	947	2413	1536	1889	339	364	137	67	841	555	1596	16888
<b>2002</b>	236	1712	298	201	483	849	1853	1044	2174	1615	1805	233	371	158	36	839	474	1625	16006
<b>2003</b>	250	1939	274	183	472	1010	1883	907	2239	1707	1640	255	336	166	35	763	511	1698	16268
<b>2004</b>	265	1744	198	131	455	1071	1938	936	2116	1577	2081	282	345	147	26	578	458	1825	16173
<b>2005</b>	254	1877	169	123	490	631	1904	996	2200	1555	2218	409	393	212	27	496	415	1976	16345
<b>2006</b>	223	2098	152	158	539	395	1820	892	2033	1668	2210	387	411	166	34	507	379	2026	16098
<b>2007</b>	225	2375	134	191	603	374	1604	953	2154	1466	2095	524	461	183	29	586	310	1919	16186
<b>2008</b>	253	2387	142	228	597	414	1546	947	1952	1423	1890	622	492	185	29	574	344	1870	15895
<b>tot</b>	3115	21421	4272	2188	5577	7131	21905	10390	25213	16956	20195	4190	4290	1687	1526	8726	5659	18616	183057

## Mobilità passiva delle SDO

La mobilità passiva extraregione non è disponibile per gli anni 2003 e 2004.

Negli archivi di mobilità passiva intra-regionale sono stati riscontrati pochi record duplicati (in totale 75 dal 1998 al 2008), che differivano soltanto per il campo "Posizione contabile"; poiché l'informazione non interessa, è stato tenuto un solo record, ovvero quello più completo nei campi "Diagnosi principale", "Tessera Sanitaria" e "Codice Fiscale", nell'ordine.

Il linkage è stato eseguito con l'anagrafe sanitaria 2010-02.

**Tab. 2.7:** Distribuzione record per anno e provenienza (extra- o intra-regionale) – mobilità passiva SDO

anno	nr record		
	extra	intra	totale
1998		47210	47210
1999		48797	48797
2000	4457	50215	54672
2001	4751	48427	53178
2002	*	46811	46811
2003		45032	45032
2004		45167	45167
2005	5460	42569	48029
2006	5439	43085	48524
2007	5815	42861	48676
2008	6087	42025	48112
<b>tot</b>	<b>32009</b>	<b>502199</b>	<b>534208</b>

\* campi anagrafici non disponibili

La distribuzione è, nel complesso, omogenea, a meno degli anni 2003 e 2004 per i quali non è disponibile la passiva extraregione e del 2002 per cui non sono disponibili i dati anagrafici con cui effettuare il linkage con l'anagrafe.

## Prescrizioni farmaceutiche

Il linkage è stato eseguito con l'anagrafe sanitaria 2010-02.

**Tab. 2.8:** Distribuzione record per anno e mese – Prescrizioni farmaceutiche

	2005	2006	2007	2008
gen	333759	357847	387532	418740
feb	324579	332726	351325	389271
mar	338234	376964	382068	384404
apr	312141	328948	351016	388379
mag	331443	371722	396095	399541
giu	313456	345990	353519	376542
luglio	295668	327785	358629	385944
ago	309493	331528	347365	341971
set	312200	330733	332666	380664
ott	306260	354559	390860	404335
nov	321010	359166	378814	369442
dic	340228	344249	354891	383671
<b>tot</b>	<b>3838471</b>	<b>4162217</b>	<b>4384780</b>	<b>4622904</b>

Nel complesso non si evidenziano aspetti problematici.

E' stata valutata la distribuzione per anno dei record per classe ATC di prescrizione farmaceutica (vedasi **Allegato B.2**).

**Tab. 2.9:** Distribuzione record per classe ATC e anno – Prescrizioni farmaceutiche

	<b>A</b>	<b>B</b>	<b>C</b>	<b>D</b>	<b>G</b>	<b>H</b>	<b>J</b>	<b>L</b>	<b>M</b>	<b>N</b>	<b>P</b>	<b>R</b>	<b>S</b>	<b>V</b>	<b>nr</b>	<b>n.d.</b>	<b>tot</b>
<b>2005</b>	419342	296253	1381809	11714	149376	141343	369887	19362	182139	252193	8468	188613	63269	539	335494	18670	<b>3838471</b>
<b>2006</b>	480115	311523	1463315	11870	153349	151910	368880	19594	198279	277125	9112	192609	67830	600	444455	11651	<b>4162217</b>
<b>2007</b>	524664	318096	1520625	12374	152217	164084	398093	18352	204722	284841	9486	197472	69891	13985	492047	3831	<b>4384780</b>
<b>2008</b>	575819	328362	1597170	13367	151563	173709	416025	18980	217523	305457	9312	197256	70920	16707	526059	4675	<b>4622904</b>
<b>tot</b>	<b>1582821</b>	<b>959276</b>	<b>4587184</b>	<b>37766</b>	<b>457860</b>	<b>490407</b>	<b>1185112</b>	<b>57040</b>	<b>621580</b>	<b>869158</b>	<b>27975</b>	<b>588322</b>	<b>208962</b>	<b>31831</b>	<b>1792195</b>		

**Note:**

Nr: ATC numerici

n.d.: ATC non definito

La distribuzione mostra complessivamente una tendenza all'aumento delle prescrizioni.

### **Fase 3: record linkage degli archivi con l'Anagrafe Sanitaria**

Tempo di realizzazione: maggio 2009 – gennaio 2010

#### *3.1 - Valutazione storicizzazione dell'archivio anagrafico*

Questa fase è stata svolta utilizzando i file di anagrafe sanitaria disponibili per gli anni precedenti al 2010, ossia i file dei mesi 2006-03,2007-01,2008-04,2009-05, confrontati con l'anagrafe sanitaria 2010-02

**Tab. 3.1:** Incrocio AS 2006,2007,2008,2009 con AS 2010

<b>Anno anagrafe</b>	<b>2006</b>	<b>2007</b>	<b>2008</b>	<b>2009</b>
<b>totale record</b>	407058	413964	425805	433649
<b>nr record linkati<sup>^</sup> con AS 2010</b>	376195	384618	399083	409777
<b>nr record non linkati<sup>^</sup></b>	30863	29346	26722	23872
<b>% record non linkati<sup>^</sup></b>	7.6%	7%	6.3%	5.5%
<b>nr record linkati<sup>^^</sup> con AS 2010</b>	382810	389822	402416	411062
<b>nr record non linkati<sup>^^</sup></b>	24248	24142	23389	22587
<b>% record non linkati<sup>^^</sup></b>	5.6%	5.8%	5.5%	5.2%
<b>nr record linkati<sup>^^^</sup> con AS 2010</b>	401759	410234	423876	432733
<b>nr record non linkati<sup>^^^</sup></b>	5299	3730	1929	916
<b>% record non linkati<sup>^^^</sup></b>	1.3%	0.9%	0.4%	0.2%

<sup>^</sup> utilizzando TS,CF, nome, cognome, data di nascita

<sup>^^</sup> utilizzando solo TS

<sup>^^^</sup> utilizzando solo CF

Per approfondire il confronto fra le due Anagrafi più recenti è stata incrociata l'AS 2006, 2007, 2008 e 2009 con l'AS 2010 per i seguenti passi di chiave, gli stessi utilizzati per l'incrocio fra Anagrafe e altre fonti:

step 1: tessera sanitaria, codice fiscale, cognome, nome, data di nascita

step 2: tessera sanitaria, cognome, nome, data di nascita  
step 3: tessera sanitaria, cognome, data di nascita  
step 4: tessera sanitaria, cognome, nome  
step 5: tessera sanitaria, nome, data di nascita  
step 6: codice fiscale, cognome, nome, data di nascita  
step 7: codice fiscale, cognome, data di nascita  
step 8: codice fiscale, cognome, nome  
step 9: codice fiscale, nome, data di nascita  
step 10: cognome, nome, data di nascita  
step 11: tessera sanitaria  
step 12: codice fiscale

Tab. 3.2: Incrocio AS 2006,2007,2008,2009 con AS 2010 – dettaglio per passo di chiave

<i>Passo di chiave</i>	2006				2007			
	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
K0	<b>407058</b>				<b>413964</b>			
<b>TCFCNDN</b>	376195	376195	92,42	92,42	384618	384618	92,91	92,91
<b>TCNDN</b>	2810	379005	0,69	93,11	1446	386064	0,35	93,26
<b>TCDN</b>	1984	380989	0,49	93,60	1832	387896	0,44	93,70
<b>TCN</b>	357	381346	0,09	93,68	332	388228	0,08	93,78
<b>TNDN</b>	1186	382532	0,29	93,97	1344	389572	0,32	94,11
<b>CFCNDN</b>	23075	405607	5,67	99,64	23040	412612	5,57	99,67
<b>CFCDN</b>	56	405663	0,01	99,66	44	412656	0,01	99,68
<b>CFCN</b>	1	405664	0,00	99,66	0	412656	0,00	99,68
<b>CFNDN</b>	59	405723	0,01	99,67	49	412705	0,01	99,70
<b>CNDN</b>	452	406175	0,11	99,78	356	413061	0,09	99,78
<b>TS</b>	<b>274</b>	<b>406449</b>	<b>0,07</b>	<b>99,85</b>	<b>248</b>	<b>413309</b>	<b>0,06</b>	<b>99,84</b>
<b>CF</b>	<b>9</b>	<b>406458</b>	<b>0,00</b>	<b>99,85</b>	<b>5</b>	<b>413314</b>	<b>0,00</b>	<b>99,84</b>
<b>NIENTE</b>	<b>600</b>	<b>407058</b>	<b>0,15</b>	<b>100</b>	<b>650</b>	<b>413964</b>	<b>0,16</b>	<b>100</b>

<i>Passo di chiave</i>	2008				2009			
	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
K0	<b>425805</b>				<b>433649</b>			
<b>TCFCNDN</b>	399083	399083	93,72	93,72	409777	409777	94,50	94,50
<b>TCNDN</b>	638	399721	0,15	93,87	439	410216	0,10	94,60
<b>TCDN</b>	1374	401095	0,32	94,20	393	410609	0,09	94,69

<b>TCN</b>	166	401261	0,04	94,24	88	410697	0,02	94,71
<b>TNDN</b>	995	402256	0,23	94,47	299	410996	0,07	94,78
<b>CFCNDN</b>	22725	424981	5,34	99,81	22437	433433	5,17	99,95
<b>CFCDN</b>	20	425001	0,00	99,81	3	433436	0,00	99,95
<b>CFCN</b>	0	425001	0,00	99,81	1	433437	0,00	99,95
<b>CFNDN</b>	30	425031	0,01	99,82	6	433443	0,00	99,95
<b>CNDN</b>	127	425158	0,03	99,85	76	433519	0,02	99,97
<b>TS</b>	160	425318	0,04	99,89	66	433585	0,02	99,99
<b>CF</b>	5	425323	0,00	99,89	0	433585	0,00	99,99
<b>NIENTE</b>	<b>482</b>	<b>425805</b>	<b>0,11</b>	<b>100,00</b>	<b>64</b>	<b>433649</b>	<b>0,01</b>	<b>100,00</b>

**Legenda:**

**T:** Tessera Sanitaria **CF:** Codice Fiscale **C:** Cognome **N:** Nome **DN:** Data di Nascita

record che nell'incrocio degli archivi con l'Anagrafe 2010 verranno comunque linkati

record eliminati da un anno all'altro

### 3.2 - Individuazione e valutazione delle chiavi di linkage utilizzabili per singolo archivio

Prima di procedere al record linkage fra archivi ed Anagrafe Sanitaria sono state individuate le variabili da utilizzare nel linkage, per impostare l'algoritmo specifico per ogni archivio; si è quindi valutata la completezza dei campi stessi.

Le chiavi di linkage individuate nel complesso sono le seguenti:

- ❖ **Tessera sanitaria (TS)**
- ❖ **Codice fiscale (CF)**
- ❖ **Nome**
- ❖ **Cognome**
- ❖ **Data di nascita**

Sono stati valutati per ogni fonte la completezza, per anno e nel totale, delle singole chiavi di linkage ed il loro formato.

## Mortalità

**Tab. 3.3:** Percentuale di completezza chiavi di linkage per anno - mortalità

chiavi	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
<b>TS</b>	campo non disponibile nel file									
<b>CF</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>nome</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>cognome</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>data nascita</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

## Schede di dimissione ospedaliera (SDO)

**Tab. 3.4:** Percentuale di completezza chiavi di linkage per anno – SDO

chiavi	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato testo	99,96	100,00	100,00	100,00	100,00	100,00	100,00	83,60	100,00	82,77	81,30
<b>CF</b> formato testo	99,68	100,00	100,00	100,00	100,00	100,00	100,00	99,20	100,00	99,53	98,16
<b>nome</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>cognome</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>data nascita</b> formato testo	100,00	100,00	100,00	100,00	99,99	100,00	100,00	100,00	100,00	100,00	100,00

## Mobilità passiva delle SDO

**Tab. 3.5:** Percentuale di completezza chiavi di linkage per anno – mobilità passiva SDO intra-regionale

chiavi	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato testo	100,00	100,00	100,00	99,98	100,00	100,00	100,00	90,68	100,00	86,22	87,47
<b>CF</b> formato testo	100,00	100,00	100,00	98,16	100,00	101,00	100,00	99,58	100,00	99,60	99,64
<b>nome</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>cognome</b> formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00
<b>data nascita</b> formato testo	99,86	99,97	99,87	99,96	99,82	99,68	100,00	100,00	100,00	100,00	100,00

**Tab. 3.6:** Percentuale di completezza chiavi di linkage per anno – mobilità passiva SDO extra-regionale

chiavi	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato testo	10,30	8,37	0,00	n.d.		8,75	7,61	6,79	6,95
<b>CF</b> formato testo	96,31	98,03	0,00	n.d.		98,85	100,99	99,12	99,63
<b>nome</b> formato testo	56,81	62,20	0,00	n.d.		66,30	67,39	57,39	56,61
<b>cognome</b> formato testo	57,05	62,20	0,00	n.d.		66,61	67,39	57,39	56,61
<b>data nascita</b> formato testo	100,00	100,00	0,00	n.d.		66,61	67,39	57,39	56,61

n.d. dato non disponibile

## Prescrizioni farmaceutiche

**Tab. 3.7:** Percentuale di completezza chiavi di linkage per anno – Prescrizioni farmaceutiche

chiavi	2005	2006	2007	2008
<b>TS</b> formato testo	Campo non disponibile			
<b>CF</b> formato testo	100,00	100,00	100,00	100,00
<b>nome</b>	Campo non disponibile			
<b>cognome</b>	Campo non disponibile			
<b>data nascita</b>	Campo non disponibile			

## Esenzioni Ticket (ET)

**Tab. 3.8:** Percentuale di completezza chiavi di linkage per anno – Esenzioni Ticket

chiavi	2000	2001	2002	2003	2004	2005	2006	2007	2008
<b>TS</b> formato testo	Campo non disponibile								
<b>CF</b> formato testo	Campo non disponibile								
<b>nome</b> formato testo	Campo non disponibile								
<b>cognome</b> formato testo	Campo non disponibile								
<b>data nascita</b> formato data/ora	Campo non disponibile								
<b>Codice assistito aziendale</b> Formato testo	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00	100,00

La procedura utilizzata per il record linkage è di tipo semi-deterministico, ovvero un record è considerato riconducibile al medesimo soggetto solo se la corrispondenza fra campi riferiti alla stessa variabile è perfetta, con una procedura a cascata di passi sempre meno restrittivi.

Le informazioni anagrafiche contenute in tutti gli archivi sono state utilizzate per identificare i soggetti nell'Anagrafe mediante i seguenti 5 passi di chiave:

- step 1: tessera sanitaria, cognome, nome, data di nascita
- step 2: tessera sanitaria, cognome, data di nascita
- step 3: tessera sanitaria, nome, data di nascita
- step 4: tessera sanitaria, cognome, nome
- step 5: cognome, nome, data di nascita

Negli archivi dove è presente anche il Codice Fiscale tra le variabili, prima dello step 5 sono stati effettuati i seguenti passi:

- step 1 (CF): codice fiscale, cognome, nome, data di nascita
- step 2 (CF): codice fiscale, cognome, data di nascita
- step 3 (CF): codice fiscale, nome, data di nascita
- step 4 (CF): codice fiscale, cognome, nome

Negli archivi dove è presente anche il Codice Fiscale tra le variabili sono stati effettuati infine i seguenti passi:

- step 1: codice fiscale, tessera sanitaria
- step 2: tessera sanitaria
- step 3: codice fiscale

Nell'archivio delle esenzioni è presente solo il codice assistito aziendale, per cui il linkage con l'anagrafe sanitaria si è basato solo su tale campo.

L'incrocio dei farmaci con l'anagrafe avviene mediante il linkage con il solo codice fiscale, essendo l'unico campo presente.

Nell'archivio della mortalità non è presente il campo "tessera sanitaria".

Volendo dare una visione d'insieme, nella seguente tabella sono evidenziate le caselle corrispondenti ai passi di chiave utilizzati nel record linkage dello specifico archivio:

**Tab. 3.9:** Passi di chiave utilizzati nel record linkage, per archivio

Passi chiave	SDO	Mortalità	MP	ET	Prescrizioni Farmaceutiche
TS C N DN					
TS C DN					
TS C N					
TS N DN					
CF C N DN					
CF C DN					
CF C N					
CF N DN					
C N DN					
TS CF					
TS					
CF					
CAA					

**Legenda:**

**TS:** Tessera Sanitaria

**C:** Cognome

**N:** Nome

**DN:** Data di Nascita

**CF:** Codice Fiscale

**CAA:** Codice assistito aziendale

*3.3 - Esecuzione di procedure di record linkage semi-deterministico a cascata*

E' stata elaborata una procedura per effettuare il record linkage mediante script sql in MSSQL 2005, con l'obiettivo di massimizzare standardizzazione (replicabilità) e sicurezza.

La procedura è stata elaborata sostanzialmente de novo.

Sono stati formulate una serie di istruzioni sql eseguite a cascata, individuando mediante un apposito campo qual è la chiave che ha prodotto il linkage. In tal modo è stato possibile individuare tutti i record agganciati.

### 3.4 - Valutazione delle percentuali di linkage

Si riportano di seguito le tabelle con il numero assoluto di record linkati per singolo passo di chiave, il numero degli appaiati cumulati nonché dei record non agganciati, con le relative percentuali.

#### Mortalità

**Tab. 3.10:** Percentuali di incrocio per singolo passo di chiave e totale - mortalità

Fonte	Periodo	Passo di chiave	Appaiati	Appaiati cumulati	% appaiati	% appaiati cumulati
mortalità	1998-2008	K0	<b>34926</b>			
		CFCNDN	28739	28739	82,29	82,29
		CFCDN	31	28770	0,09	82,37
		CFCN	5611	34381	16,07	98,44
		CFNDN	114	34495	0,33	98,77
		CNDN	407	34902	1,17	99,93
		CF	24	<b>34926</b>	0,07	100,00

\* incrocio effettuato con tutte le anagrafi disponibili

non appaiati	<b>455</b>	<b>1,29</b>
<b>Totale</b>	<b>35381</b>	

Schede di dimissione ospedaliera (SDO)

**Tab. 3.11:** Percentuali di incrocio per singolo passo di chiave e totale - SDO

<i>Fonte</i>	<i>Periodo</i>	<i>Passo di chiave</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
<b>SDO</b>	<b>1998-2008</b>	K0	<b>183058</b>			
		<b>TCNDN</b>	16655	16655	9,10	9,10
		<b>TCDN</b>	1784	18439	0,97	10,07
		<b>TCN</b>	1443	19882	0,79	10,86
		<b>TNDN</b>	998	20880	0,55	11,41
		<b>TCFCNDN</b>	118401	139281	64,68	76,09
		<b>CFCNDN</b>	14980	154261	8,18	84,27
		<b>CFCDN</b>	361	154622	0,20	84,47
		<b>CFCN</b>	6	154628	0,00	84,47
		<b>CFNDN</b>	242	154870	0,13	84,60
		<b>CNDN</b>	26914	181784	14,70	99,30
		<b>TS</b>	488	182272	0,27	99,57
		<b>CF</b>	786	183058	0,43	100,00
<b>non appaiati</b>			<b>51445</b>			<b>21,9</b>
			<b>di cui dell'ASL</b>	<b>6032</b>		<b>2,6</b>
<b>totale</b>			<b>234503</b>			

L'archivio SDO contiene tutti i ricoveri effettuati dai presidi ospedalieri dell'ASL 2 e non soltanto quelli riferiti ad assistiti della stessa.

Mobilità passiva delle SDO

Tab. 3.12: Percentuali di incrocio per singolo passo di chiave e totale – Mobilità passiva intra-regionale

<i>Fonte</i>	<i>Periodo</i>	<i>Passo di chiave</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
<b>MP intra</b>	<b>1998-2008</b>	K0	<b>502199</b>			
		TCNDN	69324	69324	13,80	13,80
		TCDN	4066	73390	0,81	14,61
		TCN	6610	80000	1,32	15,93
		TNDN	2575	82575	0,51	16,44
		TCFCNDN	323430	406005	64,40	80,85
		CFCNDN	60680	466685	12,08	92,93
		CFCDN	842	467527	0,17	93,10
		CFCN	114	467641	0,02	93,12
		CFNDN	625	468266	0,12	93,24
		CNDN	32533	500799	6,48	99,72
		TS	1349	502148	0,27	99,99
		CF	51	502199	0,01	100,00
<b>non appaiati</b>			<b>9568</b>			<b>1,87</b>
<b>Totale</b>			<b>511767</b>			

Tab. 3.13: Percentuali di incrocio per singolo passo di chiave e totale – Mobilità passiva extra-regionale

<i>Fonte</i>	<i>Periodo</i>	<i>Passo di chiave</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
<b>MP extra</b>	<b>1998-2008</b>	K0	<b>32009</b>			
		TCNDN	85	85	0,27	0,27
		TCDN	23	108	0,07	0,34

	<b>TCN</b>	7	115	0,02	0,36
	<b>TNDN</b>	3	118	0,01	0,37
	<b>TCFCNDN</b>	573	691	1,79	2,16
	<b>CFCNDN</b>	17299	17990	54,04	56,20
	<b>CFCDN</b>	447	18437	1,40	57,60
	<b>CFCN</b>	4	18441	0,01	57,61
	<b>CFNDN</b>	275	18716	0,86	58,47
	<b>CNDN</b>	878	19594	2,74	61,21
	<b>TS</b>	6	19600	0,02	61,23
	<b>CF</b>	12409	32009	38,77	100,00
<b>non appaiati totale</b>		<b>7763</b>	<b>39772</b>		<b>19,52</b>

Per quanto riguarda la mobilità passiva extra-regionale, la percentuale di assistiti non appaiati è molto elevata. In parte è dovuta alla mancata valorizzazione dei dati anagrafici nella vista dei dati del 2002.

### Prescrizioni farmaceutiche

**Tab. 3.14:** Percentuali di incrocio per singolo passo di chiave e totale – Prescrizioni farmaceutiche

<i>Fonte</i>	<i>Periodo</i>	<i>Passo di chiave</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>
farmaci	2005-2008	K0	17008372		
		CF	17008372	17008372	100
<b>non appaiati Totale</b>			<b>46617</b>	<b>17054989</b>	<b>0.2%</b>

Esenzioni Ticket (ET)

**Tab. 3.15:** Percentuali di incrocio per singolo passo di chiave e totale – Esenzioni Ticket

<i>Fonte</i>	<i>Periodo</i>	<i>Passo di chiave</i>	<i>Appaiati</i>	<i>Appaiati cumulati</i>	<i>% appaiati</i>	<i>% appaiati cumulati</i>
esenzioni	2010	K0	213917			
		CAA	213917	213917	100	100
			<b>non appaiati</b>	<b>0</b>		<b>0</b>

L'archivio delle ET è gestito in modo analogo all'Anagrafe; la quota di linkage è totale.

## Fase 4: produzione di stime di frequenza

Tempo di realizzazione: gennaio 2010 – febbraio 2010

### 4.1 - Applicazione agli archivi di algoritmi forniti

Gli algoritmi applicati alla popolazione dell'ASL 2 corrispondono a quelli presentati nel Supplemento della Rivista "Epidemiologia e Prevenzione", Anno 32 (3), maggio-giugno 2008.

### Risultati

IMA

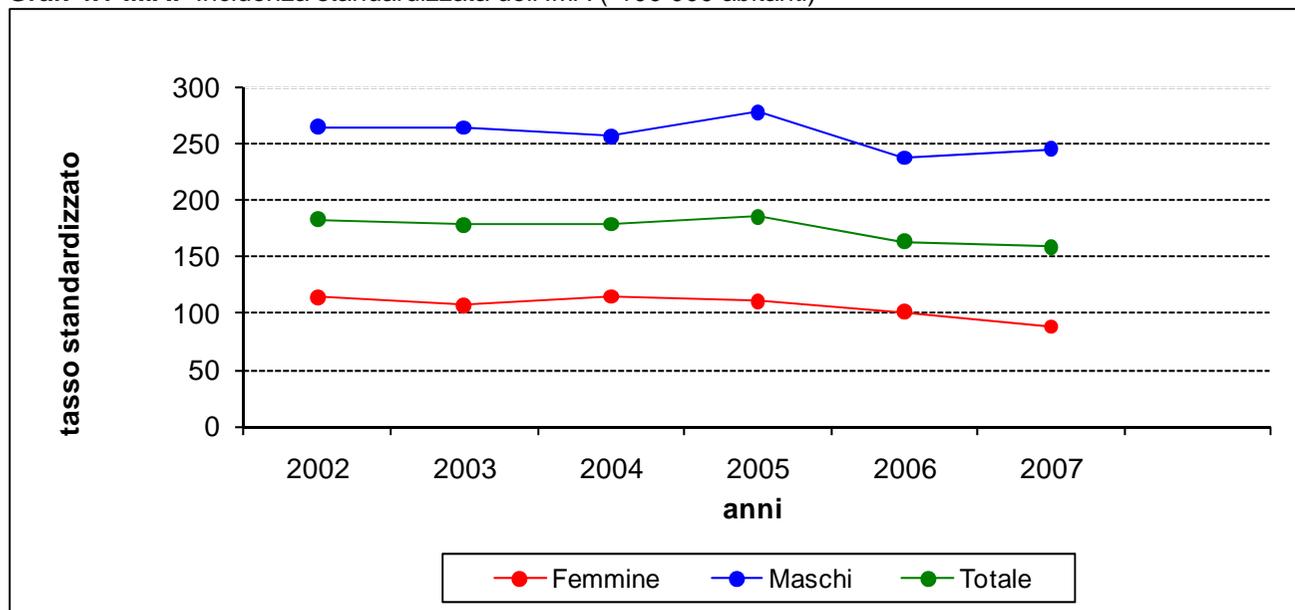
L'anno 2008 non è stato considerato in quanto non era disponibile il dato relativo alle cause di morte.

**Tab.4.1-IMA:** Incidenza grezza e standardizzata dell'IMA (\*100 000 abitanti)

anno	sezzo	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand.to
2002	F	266	170978	155,58	114,30
2002	M	425	159586	266,31	264,73
2002	totale	691,00	330564	209,04	182,94
2003	F	253	172910	146,32	107,59
2003	M	430	161409	266,40	264,13
2003	totale	683,00	334319	204,30	178,16
2004	F	273	176516	154,66	114,77
2004	M	426	164460	259,03	256,79
2004	totale	699,00	340976	205,00	178,95
2005	F	280	179949	155,60	110,87
2005	M	472	167921	281,08	278,07

2005	totale	752,00	347870	216,17	185,39
2006	F	259	183889	140,85	101,29
2006	M	413	171926	240,22	237,47
2006	totale	672,00	355815	188,86	163,78
2007	F	234	184413	126,89	88,41
2007	M	430	172595	249,14	245,62
2007	totale	664,00	357008	185,99	159,00

**Graf. 4.1-IMA:** Incidenza standardizzata dell'IMA (\*100 000 abitanti)



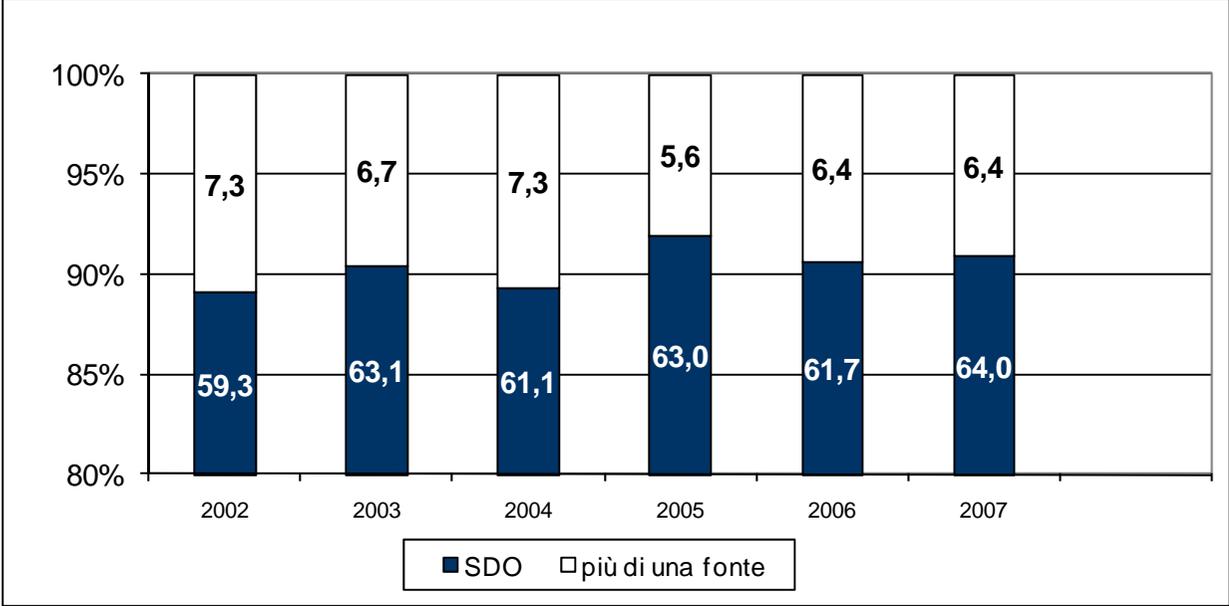
Tab. 4.2-IMA: Distribuzione casi per classi d'età e anno

Classi età	anno						Totale
	2002	2003	2004	2005	2006	2007	
0-14	0	0	0	0	0	0	0
	0	0	0	0	0	0	
15-34	0	0	0	0	0	0	0
	0	0	0	0	0	0	
35-54	72	81	69	65	72	78	492
	10,42	11,86	9,87	8,64	10,71	11,75	
55-64	103	99	113	121	103	95	731
	14,91	14,49	16,17	16,09	15,33	14,31	
65-74	148	175	151	158	146	139	1033
	21,42	25,62	21,60	21,01	21,73	20,93	
75+	368	328	366	408	351	352	2398
	53,26	48,02	52,36	54,26	52,23	53,01	
<b>Totale</b>	691	683	699	752	672	664	4654

Tab. 4.3-IMA: Contributo esclusivo delle fonti per anno di stima

Contributo fonti												
	2002		2003		2004		2005		2006		2007	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
<b>SDO</b>	480	59,33	498	63,12	500	61,05	533	63,00	475	61,69	488	64,04
<b>CM</b>	270	33,37	238	30,16	259	31,62	266	31,44	246	31,95	225	29,53
<b>più di una fonte</b>	59	7,29	53	6,72	60	7,33	47	5,56	49	6,36	49	6,43
<b>Totale</b>	809	100,00	789	100	819	100,00	846	100,00	770	100	762	100,00

**Graf. 4.2-IMA:** Contributo esclusivo delle fonti per anno di stima

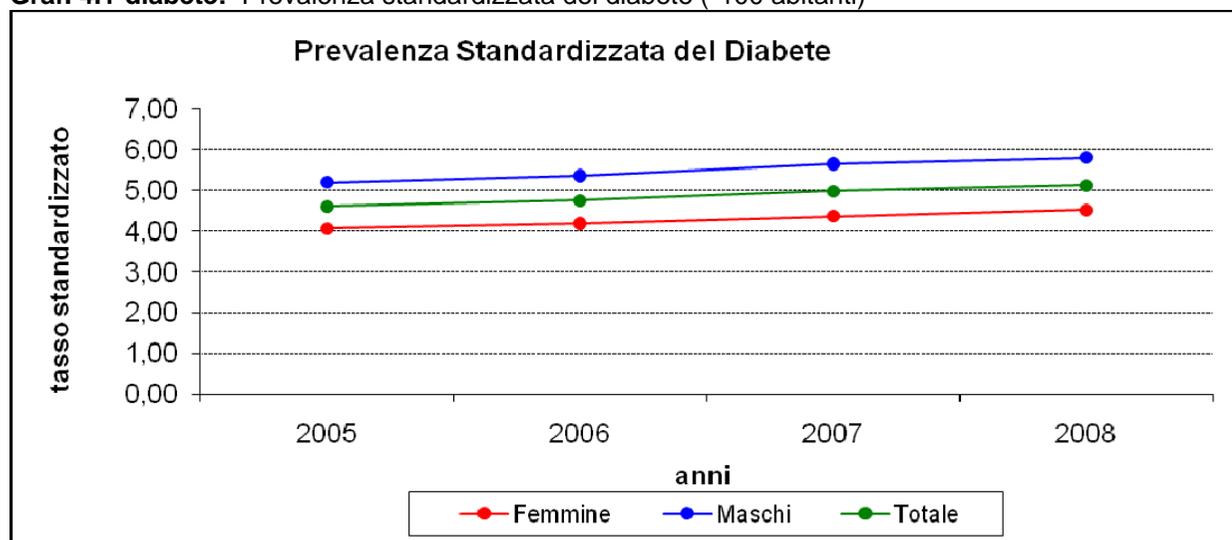


## Diabete

**Tab. 4.1-diabete:** Prevalenza grezza e standardizzata del diabete (\*100 abitanti)

anno	sesso	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand.to
2005	F	8955	179949	4,98	<b>4,06</b>
2005	M	8957	167921	5,33	<b>5,20</b>
2005	totale	17912	347870	5,15	<b>4,61</b>
2006	F	9327	183889	5,07	<b>4,19</b>
2006	M	9405	171926	5,47	<b>5,36</b>
2006	totale	18732	355815	5,26	<b>4,75</b>
2007	F	9831	184413	5,33	<b>4,37</b>
2007	M	9999	172595	5,79	<b>5,65</b>
2007	totale	19830	357008	5,55	<b>4,98</b>
2008	F	10291	187649	5,48	<b>4,52</b>
2008	M	10413	175266	5,94	<b>5,80</b>
2008	totale	20704	362915	5,70	<b>5,13</b>

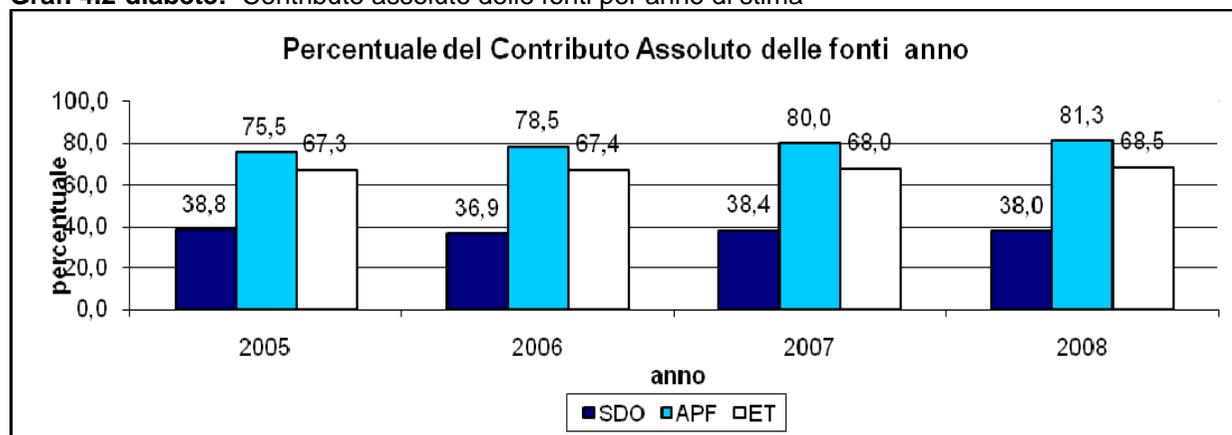
Graf. 4.1-diabete: Prevalenza standardizzata del diabete (\*100 abitanti)



**Tab. 4.2-diabete:** Distribuzione casi per classi d'età e anno di stima

<b>Distribuzione per classi di età ed anno</b>					
<b>Classi età</b>	<b>anno</b>				<b>Totale</b>
<b>Frequenza Percent col</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>	<b>2008</b>	
<b>0-14</b>	55	58	53	48	214
	0,31	0,31	0,27	0,23	
<b>15-34</b>	339	349	385	428	1501
	1,89	1,86	1,94	2,07	
<b>35-54</b>	2208	2331	2492	2588	9619
	12,33	12,44	12,57	12,50	
<b>55-64</b>	3660	3857	4015	4117	15649
	20,43	20,59	20,25	19,89	
<b>65-74</b>	5542	5658	5819	6044	23063
	30,94	30,20	29,34	29,19	
<b>75+</b>	6108	6479	7066	7479	27132
	34,10	34,59	35,63	36,12	
<b>Totale</b>	17912	18732	19830	20704	77178

**Graf. 4.2-diabete:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3-diabete:** Contributo esclusivo delle fonti, anni di stima 2004-2008

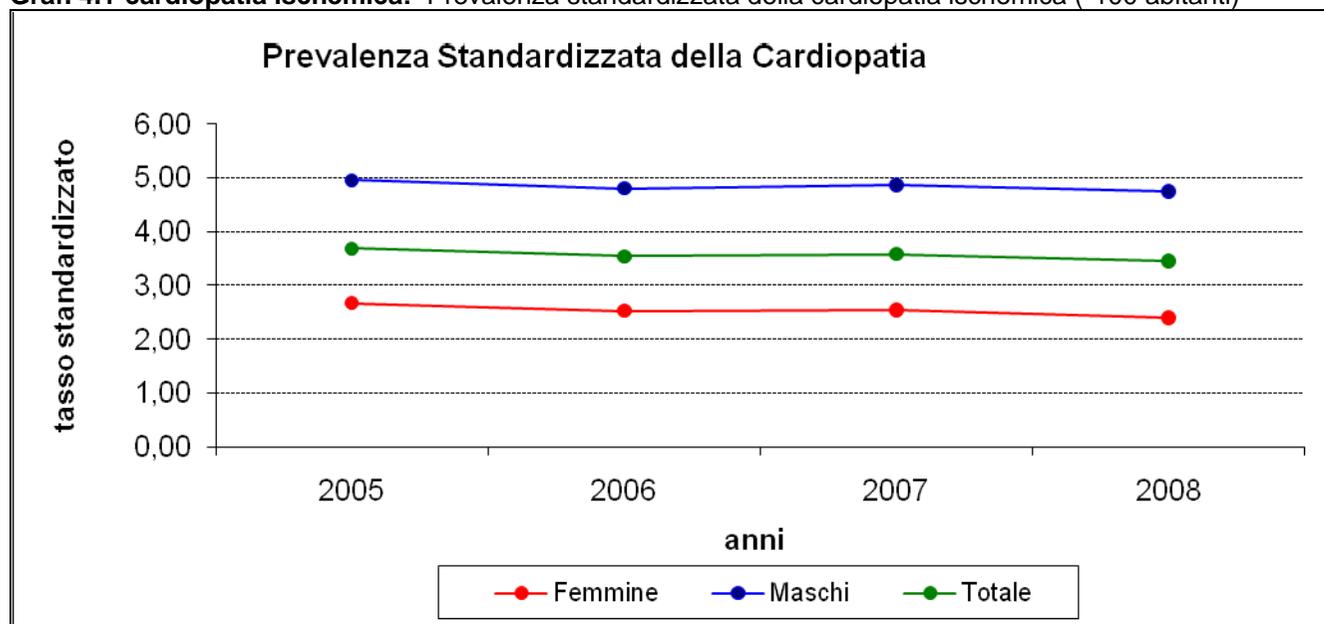
Contributo esclusivo fonti								
	2005		2006		2007		2008	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
<b>SDO</b>	2648	9,11	2304	7,8	2165	7,2	1882	7,6
<b>APF</b>	4235	14,57	4342	14,7	4306	14,3	4452	17,9
<b>ET</b>	11152	38,36	11099	37,5	10672	35,4	4756	19,1
<b>più di una fonte</b>	11037	37,96	11884	40,1	13041	43,2	13826	55,5

## Cardiopatia ischemica

**Tab. 4.1-cardiopatia ischemica:** Prevalenza grezza e standardizzata della cardiopatia ischemica (\*100 abitanti)

anno	sezzo	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand.to
2005	F	6751,00	179949,00	3,75	<b>2,67</b>
2005	M	8364,00	167921,00	4,98	<b>4,95</b>
2005	totale	15115,00	347870,00	4,35	<b>3,68</b>
2006	F	6526,00	183889,00	3,55	<b>2,53</b>
2006	M	8286,00	171926,00	4,82	<b>4,80</b>
2006	totale	14812,00	355815,00	4,16	<b>3,54</b>
2007	F	6731,00	184413,00	3,65	<b>2,54</b>
2007	M	8557,00	172595,00	4,96	<b>4,86</b>
2007	totale	15288,00	357008,00	4,28	<b>3,58</b>
2008	F	6468,00	187649,00	3,45	<b>2,40</b>
2008	M	8495,00	175266,00	4,85	<b>4,75</b>
2008	totale	14963,00	362915,00	4,12	<b>3,45</b>

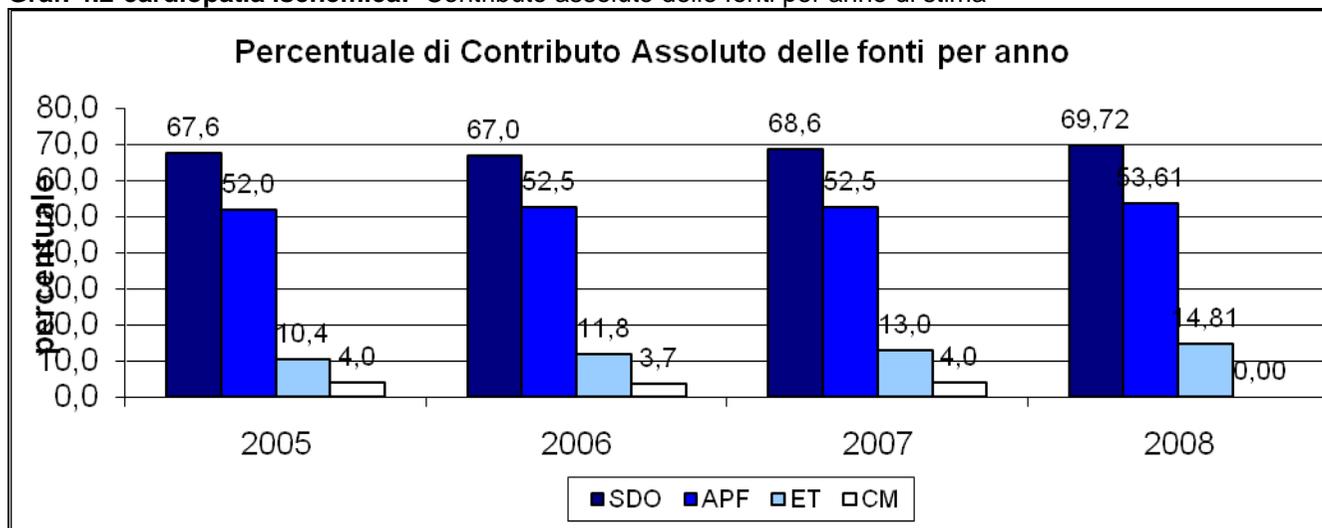
**Graf. 4.1-cardiopatia ischemica:** Prevalenza standardizzata della cardiopatia ischemica (\*100 abitanti)



**Tab. 4.2-cardiopatia ischemica:** Distribuzione casi per classi d'età e anno di stima

<b>Distribuzione per classi di età ed anno</b>					
<b>Classi età</b>	<b>anno</b>				
<b>Frequenza Percent col</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>	<b>2008</b>	<b>totale</b>
<b>0-14</b>	4	4	8	7	23
	0,03	0,03	0,05	0,05	
<b>15-34</b>	16	23	25	24	88
	0,11	0,16	0,16	0,16	
<b>35-54</b>	727	720	737	738	2922
	4,81	4,86	4,82	4,93	
<b>55-64</b>	1754	1766	1800	1794	7114
	11,60	11,92	11,77	11,99	
<b>65-74</b>	3756	3588	3584	3463	14391
	24,85	24,22	23,44	23,14	
<b>75+</b>	8858	8711	9134	8937	35640
	58,60	58,81	59,75	59,73	
<b>Totale</b>	15115	14812	15288	14963	60178

**Graf. 4.2-cardiopatia ischemica:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- cardiopatia ischemica:** Contributo esclusivo delle fonti, anni di stima 2004-2008

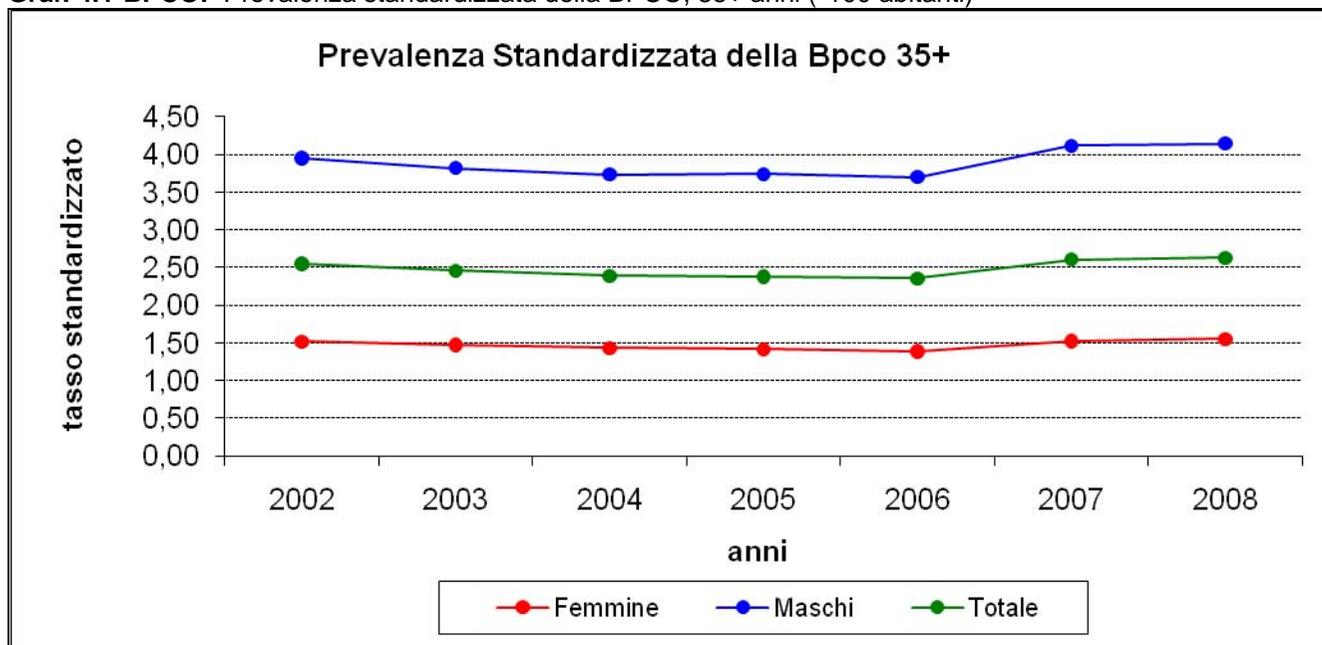
Contributo esclusivo fonti								
	2005		2006		2007		2008	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
<b>SDO</b>	8731	40,757	7740	38,128	7466	36,2	6852	35,9
<b>APF</b>	5991	27,967	5749	28,32	5588	27,1	5390	28,2
<b>ET</b>	1356	6,3299	1499	7,3842	1686	8,2	1665	8,7
<b>CM</b>	603	2,8149	554	2,7291	619	3,0	0	0,0
<b>più di una fonte</b>	4741	22,131	4758	23,438	5251	25,5	5196	27,2
<b>totale</b>	21422	100	20300	100	20610	100	19103	100,0

## BPCO

**Tab. 4.1-BPCO:** Prevalenza grezza e standardizzata della BPCO, 35+ anni (\*100 abitanti)

anno	sezzo	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand.to
2002	F	2003	109966	1,82	<b>1,52</b>
2002	M	3743	96574	3,88	<b>3,95</b>
2002	totale	5746	206540	2,78	<b>2,55</b>
2003	F	2002	111918	1,79	<b>1,47</b>
2003	M	3702	98197	3,77	<b>3,82</b>
2003	totale	5704	210115	2,71	<b>2,46</b>
2004	F	1998	114711	1,74	<b>1,43</b>
2004	M	3684	100183	3,68	<b>3,73</b>
2004	totale	5682	214894	2,64	<b>2,39</b>
2005	F	2034	117309	1,73	<b>1,42</b>
2005	M	3753	102580	3,66	<b>3,74</b>
2005	totale	5787	219889	2,63	<b>2,38</b>
2006	F	2023	119530	1,69	<b>1,39</b>
2006	M	3817	105231	3,63	<b>3,70</b>
2006	totale	5840	224761	2,60	<b>2,35</b>
2007	F	2284	120451	1,90	<b>1,52</b>
2007	M	4317	106184	4,07	<b>4,11</b>
2007	totale	6601	226635	2,91	<b>2,61</b>
2008	F	2383	122727	1,94	<b>1,55</b>
2008	M	4418	107980	4,09	<b>4,14</b>
2008	totale	6801,00	230707	2,95	<b>2,63</b>

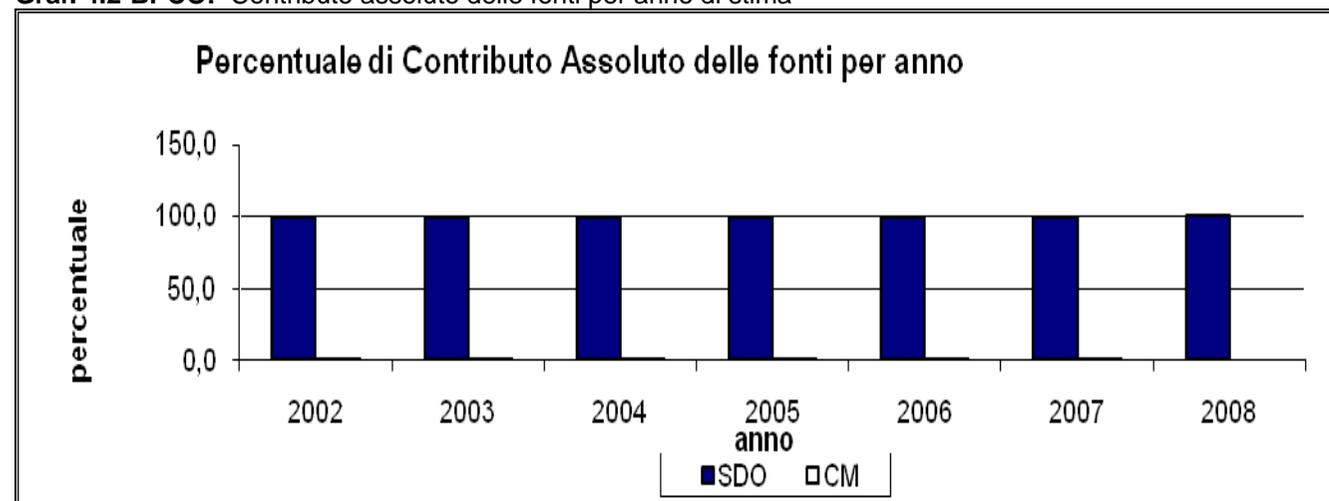
Graf. 4.1-BPCO: Prevalenza standardizzata della BPCO, 35+ anni (\*100 abitanti)



**Tab. 4.2-BPCO:** Distribuzione casi per classi d'età e anno di stima

Distribuzione per classi di età ed anno								
Classi età	anno							
Frequenza Percent col	2002	2003	2004	2005	2006	2007	2008	Totale
35-54	368	319	304	303	302	330	339	2265
	0,64	0,56	0,54	0,52	0,52	0,50	0,50	
55-64	721	690	644	613	635	662	649	4614
	12,55	12,10	11,33	10,59	10,87	10,03	9,54	
65-74	1573	1565	1521	1546	1502	1621	1608	10936
	27,38	27,44	26,77	26,72	25,72	24,56	23,64	
75+	3084	3130	3213	3325	3401	3988	4205	24346
	53,67	54,87	56,55	57,46	58,24	60,42	61,83	
<b>Totale</b>	<b>5746</b>	<b>5704</b>	<b>5682</b>	<b>5787</b>	<b>5840</b>	<b>6601</b>	<b>6801</b>	<b>42161</b>

**Graf. 4.2-BPCO:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- BPCO:** Contributo esclusivo delle fonti, anni di stima 2004-2008

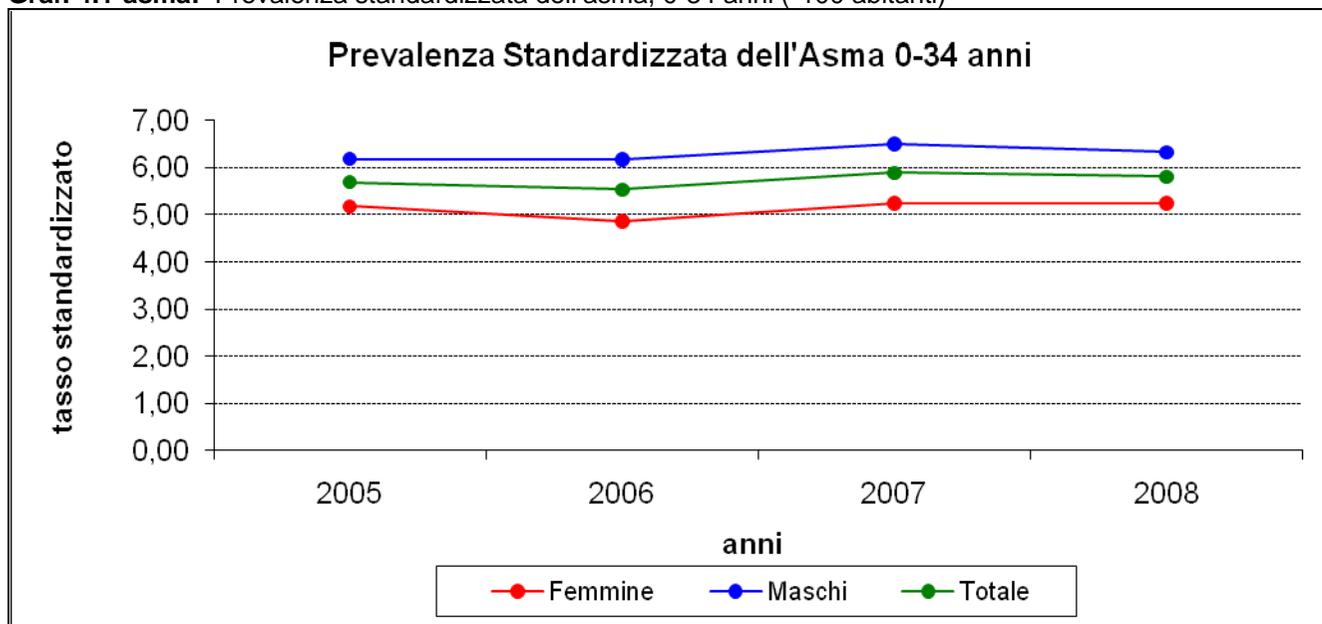
Contributo esclusivo fonti - 35 e più														
	2002		2003		2004		2005		2006		2007		2008	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
<b>SDO</b>	5689	98,7	5655	99,0	5602	98,4	5697	98,1	5762	98,5	6538	97,9	6801	100,0
<b>CM</b>	64	1,1	51	0,9	84	1,5	94	1,6	78	1,3	63	0,9	0	0,0
<b>più di una fonte</b>	10	0,2	7	0,1	7	0,1	14	0,2	7	0,1	76	1,1	0	0,0
<b>totale</b>	5763	100,0	5713	100,0	5693	100,0	5805	100,0	5847	100	6677	100	6801	100

Asma

**Tab. 4.1-asma:** Prevalenza grezza e standardizzata dell'asma, 0-34 anni (\*100 abitanti)

anno	sesso	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand.to
2005	F	3305	62640	5,28	<b>5,18</b>
2005	M	4133	65341	6,33	<b>6,19</b>
2005	totale	7438	127981	5,81	<b>5,69</b>
2006	F	3170	64359	4,93	<b>4,86</b>
2006	M	4221	66695	6,33	<b>6,17</b>
2006	totale	7391	131054	5,64	<b>5,53</b>
2007	F	3426	63962	5,36	<b>5,24</b>
2007	M	4477	66411	6,74	<b>6,50</b>
2007	totale	7903	130373	6,06	<b>5,89</b>
2008	F	3516	64922	5,42	<b>5,25</b>
2008	M	4463	67286	6,63	<b>6,33</b>
2008	totale	7979	132208,00	6,04	<b>5,81</b>

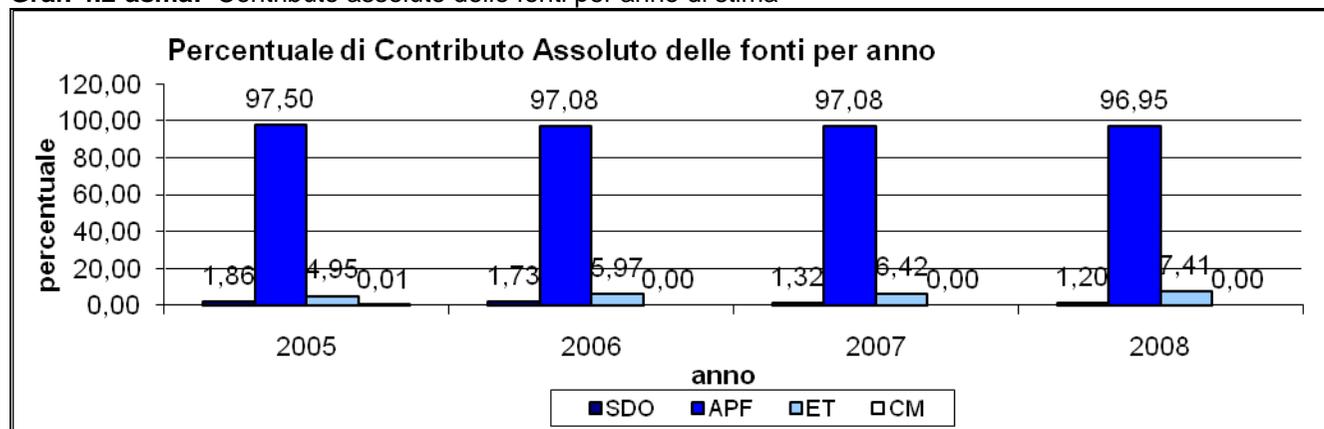
**Graf. 4.1-asma:** Prevalenza standardizzata dell'asma, 0-34 anni (\*100 abitanti)



**Tab. 4.2-asma:** Distribuzione casi per classi d'età e anno di stima

Distribuzione per classi di età ed anno					
Classi età	anno				
	2005	2006	2007	2008	Totale
0-14	4873	4651	5099	5217	19840
	65,51	62,93	64,52	65,38	
15-34	2565	2740	2804	2762	10871
	34,49	37,07	35,48	34,62	
<b>Totale</b>	<b>7438</b>	<b>7391</b>	<b>7903</b>	<b>7979</b>	<b>30711</b>

**Graf. 4.2-asma:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- asma:** Contributo esclusivo delle fonti, anni di stima 2004-2008

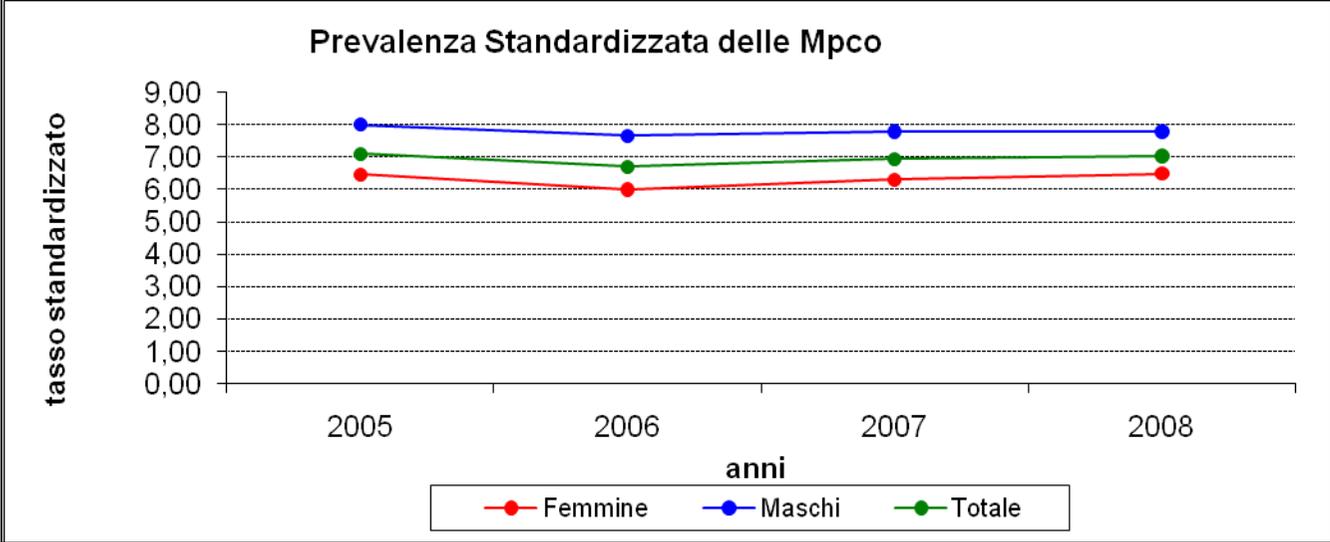
Contributo esclusivo fonti 0-34 anni								
	2005		2006		2007		2008	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
<b>SDO</b>	50	0,67	44	0,6	30	0,38	26	0,3
<b>APF</b>	6937	93,25	6833	92,5	7300	92,4	7299	91,5
<b>ET</b>	135	1,81	171	2,3	201	2,5	217	2,7
<b>CM</b>	1	0,01	0	0,0	0	0,0	0	0,0
<b>più di una fonte</b>	316	4,25	343	4,6	372	4,7	437	5,5
<b>totale</b>	7439	100,00	7391	100,0	7903	100	7979	100,0

## MPCO

**Tab. 4.1-MPCO:** Prevalenza grezza e standardizzata delle MPCO (\*100 abitanti)

anno	sezzo	numeratore tasso grezzo	denominatore tasso grezzo	tasso grezzo	tasso stand.to
2005	F	12657	179949	7,03	<b>6,46</b>
2005	M	13504	167921	8,04	<b>8,01</b>
2005	totale	26161	347870	7,52	<b>7,10</b>
2006	F	11914	183889	6,48	<b>5,99</b>
2006	M	13203	171926	7,68	<b>7,65</b>
2006	totale	25117	355815	7,06	<b>6,70</b>
2007	F	12557	184413	6,81	<b>6,30</b>
2007	M	13629	172595	7,90	<b>7,79</b>
2007	totale	26186	357008	7,33	<b>6,94</b>
2008	F	13195	187649	7,03	<b>6,49</b>
2008	M	13871	175266	7,91	<b>7,78</b>
2008	totale	27066	362915	7,46	<b>7,03</b>

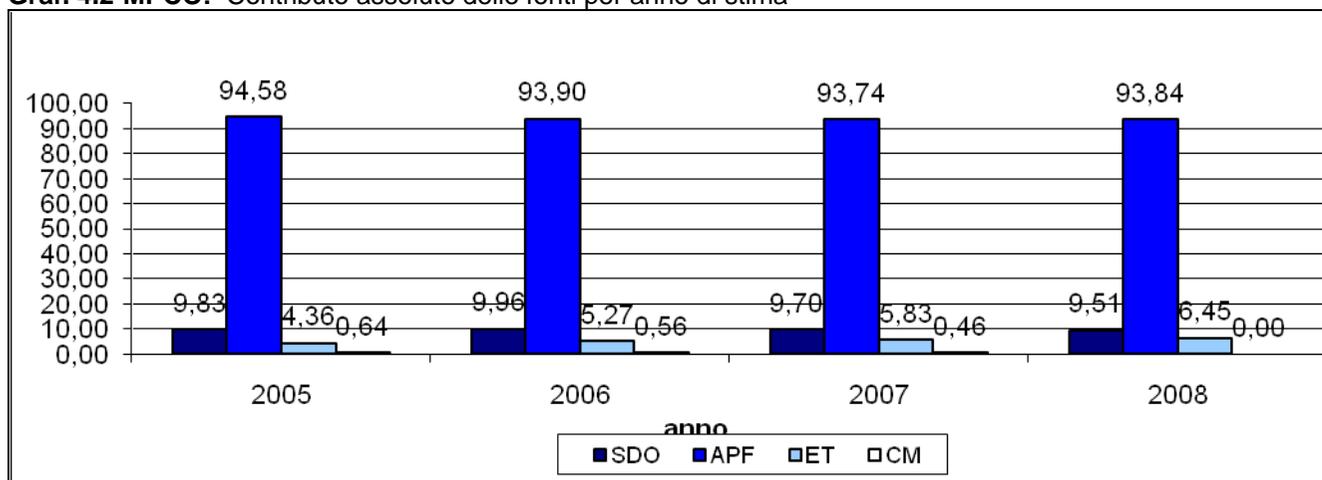
Graf. 4.1-MPCO: Prevalenza standardizzata delle MPCO (\*100 abitanti)



**Tab. 4.2-MPCO:** Distribuzione casi per classi d'età e anno di stima

<b>Distribuzione per classi di età ed anno</b>					
<b>Classi età</b>	<b>anno</b>				
<b>Frequenza Percent col</b>	<b>2005</b>	<b>2006</b>	<b>2007</b>	<b>2008</b>	<b>totale</b>
<b>0-14</b>	4917	4688	5120	5227	19952
	18,80	18,66	19,55	19,31	
<b>15-34</b>	2664	2826	2887	2832	11209
	10,18	11,25	11,02	10,46	
<b>35-54</b>	3925	4005	4336	4652	16918
	15,00	15,95	16,56	17,19	
<b>55-64</b>	2724	2545	2626	2838	10733
	10,41	10,13	10,03	10,49	
<b>65-74</b>	4593	4090	4175	4259	17117
	17,56	16,28	15,94	15,74	
<b>75+</b>	7338	6963	7042	7258	28601
	28,05	27,72	26,89	26,82	
<b>Totale</b>	26161	25117	26186	27066	104530

**Graf. 4.2-MPCO:** Contributo assoluto delle fonti per anno di stima



**Tab. 4.3- MPCO:** Contributo esclusivo delle fonti, anni di stima 2004-2008

Contributo esclusivo fonti di tutte le età								
	2005		2006		2007		2008	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%
<b>SDO</b>	1006	3,8	1059	4,2	1104	4,2	1107	4,1
<b>APF</b>	22445	85,6	21305	84,7	22123	84,4	22827	84,3
<b>ET</b>	345	1,3	412	1,6	484	1,8	550	2,0
<b>CM</b>	95	0,4	77	0,3	63	0,2	0	0,0
<b>più di una fonte</b>	2324	8,9	2306	9,2	2450	9,3	2582	9,5
<b>totale</b>	26215	100	25159	100,0	26224	100,0	27066	100,0

#### *4.2 - Confronto con stime ottenute applicando gli stessi algoritmi in altre realtà italiane*

Le altre realtà italiane prese come riferimento per il confronto delle stime di incidenza e prevalenza delle patologie di cui sopra sono quelle presenti nel Supplemento della Rivista "Epidemiologia e Prevenzione" sopra citato. Le stime pubblicate in questo documento sono state prodotte per gli anni 2002, 2003 e 2004. Le seguenti considerazioni si riferiscono a questo periodo temporale. La realtà più vicina alla ASL 2 dell'Umbria (Perugia) che abbia utilizzato questi algoritmi è l'Azienda Sanitaria 10 di Firenze.

Per quanto riguarda l'**IMA**, le stime di incidenza ottenute a Perugia sono più elevate rispetto a quelle della popolazione fiorentina tra i maschi mentre sono sostanzialmente simili per il sesso femminile. Il contributo della fonte Mortalità nella definizione della casistica è superiore nella ASL di Perugia rispetto a quella di Firenze.

La prevalenza stimata di **diabete** risulta più elevata a Perugia, in entrambi i sessi, rispetto a Firenze. Il contributo assoluto della fonte SDO è più basso; quello della farmaceutica è sovrapponibile, mentre la fonte ET per l'individuazione dei casi presenta valori sempre più elevati a Perugia. I soggetti identificati con più fonti contemporanee è cresciuto nel quadriennio ed è più alto rispetto all'anno 2003 di Firenze.

I tassi stimati a Perugia per la **cardiopatía ischemica** sono lievemente superiori a quelli di Firenze e degli altri centri. Rispetto all'anno 2003 del centro di Firenze, il contributo assoluto delle SDO è sovrapponibile, mentre quello della farmaceutica è più basso; la fonte ET, invece, è quasi doppia. Il contributo esclusivo di più fonti è in crescita dal 2005 al 2008 ma, comunque, inferiore a quello di Firenze.

A Perugia i tassi di prevalenza di **BPCO** negli uomini sono superiori di un punto percentuale rispetto a Firenze, mentre sono di poco più elevati nelle donne. La distribuzione dei contributi per fonte è comparabile a quella degli altri centri.

La prevalenza stimata di **asma** è più elevata a Perugia rispetto al centro di Firenze. I contributi esclusivi all'individuazione dei casi sono in linea per tutte le fonti tranne per l'ET, più elevata a Perugia.

Anche le stime di prevalenza di **MPCO** sono sensibilmente più elevate a Perugia in entrambi i sessi, con andamento piuttosto stabile nel periodo. La distribuzione dei contributi è in linea con gli altri centri ma lievemente più alta per la fonte ET.

## Appendice

### Allegato B.1– Gruppi di cause di morte, classificazione ICD-IX

GRUPPO	CODICI	DENOMINAZIONE
I	001-139	Malattie infettive e parassitarie
II	140-239	Tumori
III	240-279	Malattie endocrine
IV	280-289	Malattie del sangue e degli organi ematopoietici
V	290-319	Disturbi psichici
VI	320-389	Malattie del sistema nervoso e degli organi dei sensi
VII	390-459	Malattie del sistema circolatorio
VIII	460-519	Malattie dell'apparato respiratorio
IX	520-579	Malattie dell'apparato digerente
X	580-629	Malattie dell'apparato genito-urinario
XI	630-679	Complicazione della gravidanza, del parto e del puerperio
XII	680-709	Malattie della pelle e del tessuto sottocutaneo
XIII	710-739	Malattie del sistema osteomuscolare e del tessuto connettivo
XIV	740-759	Malformazioni congenite
XV	760-779	Condizioni morbose perinatali
XVI	780-799	Sintomi e stati morbosi maldefiniti
XVII	800-999	Accidenti, avvelenamenti e traumatismi

### Allegato B.2– Categorie farmaceutiche, classificazione ATC

CATEGORIA	DESCRIZIONE
A	Apparato gastrointestinale e metabolismo
B	Sangue e organi emopoietici
C	Sistema cardiovascolare
D	Dermatologici
G	Sistema genito-urinario ed ormoni sessuali
H	Preparati ormonali sistemici, esclusi gli ormoni sessuali
J	Antimicrobici generali per uso sistemico
L	Farmaci antineoplastici ed immunomodulatori
M	Sistema muscolo-scheletrico

N Sistema nervoso  
P Farmaci antiparassitari, insetticidi e repellenti  
R Sistema respiratorio  
S Organi di senso  
V Vari

## **OBIETTIVO SPECIFICO 4**

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

Grazie anche all'esperienza positiva del gruppo nazionale AIE-SISMEC e di altre esperienze qualificate, come quelle di farmaco-epidemiologia, che sono attive in Italia nel campo dell'utilizzo epidemiologico di archivi sanitari elettronici, è stato possibile riunire un gruppo di esperti .

Gli scopi, la discussione ed il programma di lavoro ha portato il gruppo a dividere il lavoro in tre aree che sono riprese qui di seguito con i risultati ottenuti in ciascuna area.

Si ritiene di vitale importanza che il programma di replicazione di sistemi integrati che costituisce la parte centrale del progetto finanziato dal CCM venga sostenuto dall'attività di elaborazione e proposta dei metodi di un gruppo di esperti.

Sono stato proposti e discussi 3 principali temi o aree su cui concentrare ed organizzare il lavoro:

1. Strumenti e metodi per l'attivazione di database integrati presso le USL;
2. sviluppo ed applicazione di algoritmi per individuare soggetti affetti da malattia nella popolazione;
3. validazione della casistica.

## COMPOSIZIONE E AFFILIAZIONI GRUPPO 'ESPERTI':

### **Lorenzo Simonato**

(Dipartimento di Medicina Ambientale e Sanità Pubblica, Università di Padova)

### **Cristina Canova**

(Dipartimento di Medicina Ambientale e Sanità Pubblica, Università di Padova)

### **Vincenzo Baldo**

(Dipartimento di Medicina Ambientale e Sanità Pubblica, Università di Padova)

### **Paolo Rigobello**

(Dipartimento di Medicina Ambientale e Sanità Pubblica, Università di Padova)

### **Alice Vianello**

(Dipartimento di Medicina Ambientale e Sanità Pubblica, Università di Padova)

### **Silvana Zanon**

(Dipartimento di Medicina Ambientale e Sanità Pubblica, Università di Padova)

### **Piergiorgio Duca**

(Unità di Statistica Medica e Biometria - Dipartimento Scienze Cliniche  
Ospedale Luigi Sacco, Milano)

### **Pietro Barbieri**

(Azienda ospedaliera di Melegnano)

### **Mariangela Vigotti**

(Dipartimento di Biologia-Università di Pisa)

### **Roberto Gnani**

(Servizio di Epidemiologia ASLTO3-Regione Piemonte)

### **Roberta Picariello**

(Servizio di Epidemiologia ASLTO3-Regione Piemonte)

### **Claudia Galassi**

(CPO Piemonte e Università Torino)

### **Annunziata Faustini**

(Dipartimento di Epidemiologia, ASL RME, Roma)

### **Silvia Cascini**

(Dipartimento di Epidemiologia, ASL RME, Roma)

### **Massimo Arcà**

(Dipartimento di Epidemiologia, ASL RME, Roma)

**Alessandro Barchielli**

(Unità di Epidemiologia, Azienda Sanitaria 10 Firenze)

**Susanna Conti**

(Ufficio di Statistica, CNESPS, Istituto. Superiore di Sanità)

**Bonora Karin**

(ASL Modena)

**De Girolamo Gianfranco**

(ASL Modena)

**Igino Fusco Moffa**

(ASL Perugia)

**Pierantonio Romor**

(INSIEL-Friuli Venezia Giulia)

**Loris Zanier**

(FVG)

#### **RIUNIONI EFFETTUATE DAL GRUPPO 'ESPERTI':**

- **11 Settembre 2008, Venezia, Polo didattico Ospedale Civile S.S. Giovanni e Paolo**
- **15-16 Gennaio 2009, Firenze, Istituto Agronomico dell'oltremare**
- **15-16 Giugno 2009, Firenze, Istituto Agronomico dell'oltremare**
- **24 Novembre 2009, Firenze, Centro studi CISL Fiesole**
- **24-25 Febbraio 2010, Venezia, Kosher Hotel Giardino del Ghetto**

## **RELAZIONE GRUPPO DI LAVORO AREA 1: Strumenti e metodi per l'attivazione di database integrati presso le USL**

### **Generalità**

L'ordinamento corrente attribuisce alle aziende USL la funzione di tutela della salute della popolazione residente. Presupposto all'esercizio di questa funzione è una buona conoscenza di come si distribuiscono, nella popolazione da tutelare, lo stato di salute e i suoi determinanti, siano essi esterni o interni al sistema sanitario.

Poiché solo per alcune patologie è possibile mettere in piedi sistemi dedicati di reperimento, registrazione e validazione di informazioni cliniche ed anamnestiche che puntano alla copertura totale (e.g. registri tumori), è ormai opinione condivisa che i dati dei sistemi informativi correnti, che si originano in occasione dei contatti tra le persone e i servizi sanitari e ai quali ci si riferisce frequentemente con il termine di "dati amministrativi", rappresentino un patrimonio di conoscenza non surrogabile.

Ovviamente, l'accuratezza dei dati amministrativi, e quindi il loro valore a fini di descrizione epidemiologica, è funzione delle caratteristiche dei sistemi di rilevazione e controllo dei dati stessi, nonché degli strumenti posti a supporto e presidio della loro qualità.

A questo vanno necessariamente ad aggiungersi il se e il come le aziende USL vengono messe nelle condizioni di accedere alla totalità dei dati amministrativi riferibili ai propri residenti e, soprattutto, la possibilità di attribuire alla stessa persona informazioni generate in momenti diversi (integrazione tra basi di dati).

Per cercare di delineare un quadro più chiaro, seppur non esaustivo, delle caratteristiche (contenuti, formati qualità, tempestività, integrabilità) del patrimonio informativo che nelle diverse regioni italiane viene reso disponibile alle aziende USL, il gruppo di esperti ha ritenuto utile predisporre un questionario ed inviarlo, con preghiera di compilazione, ai livelli regionali interessati. La rilevazione ha avuto luogo nella seconda metà del 2009 ed ha prodotto i risultati che vengono riportati e commentati nell'appendice posta al termine di questa sezione (**Allegato C**).

### **Identificazione delle fonti informative e acquisizione dei dati**

I dati amministrativi vengono generati in luoghi diversi, in occasione di eventi diversi e per finalità diverse.

I sistemi informativi "delle prestazioni" sono alimentati in primo luogo dai soggetti erogatori (ospedali, ambulatori, laboratori di analisi) che, nella maggioranza dei casi, utilizzano strumenti informatizzati primariamente orientati a supportare l'erogazione dell'assistenza. Nella maggior parte dei casi questi stessi strumenti sono progettati in modo da soddisfare, tra gli altri, il "debito informativo" previsto dalle normative regionali (esempio tipico: i sistemi ADT e la produzione dei dati SDO)

In molte situazioni, e per molti dei dati considerati, le ASL rappresentano uno snodo importante del sistema informativo: acquisiscono i dati dai presidi a gestione diretta e dalle strutture private accreditate presenti sul territorio aziendale e li inoltrano, eventualmente dopo opportune procedure di accettazione e controllo, al livello regionale.

Questo diffuso modello organizzativo è stato negli ultimi anni messo in discussione e, in qualche modo, rovesciato dall'implementazione del sistema TS da parte del Ministero dell'Economia e delle Finanze. Questo sistema, in ottemperanza a quanto stabilito dall'art.50 della L.326/2003, prevede che, per quanto riguarda le prestazioni di specialistica ambulatoriale e le prescrizioni

farmaceutiche, le strutture erogatrici e le farmacie trasmettano i dati analitici direttamente al livello nazionale. Alle Regioni è stato tuttavia consentito (comma 11 del citato art.50) di esercitare un'opzione che prevede la raccolta dei dati a livello regionale e, da qui, l'invio al sistema TS.

Per quanto riguarda l'accessibilità alle informazioni da parte delle aziende USL, il sistema TS prevede che operatori appositamente autorizzati possano:

- Accedere ad una serie di indicatori sintetici (il cosiddetto cruscotto integrato), alcuni dei quali fanno riferimento alla popolazione assistita dai MMG e PLS convenzionati con la ASL
- Scaricare i dati analitici (un file anonimo e un file di codici fiscali linkabili tra loro) relativi alle prescrizioni fruite presso le strutture di propria competenza.

La rilevazione effettuata nell'ambito di questo progetto, seppure non esaustiva, documenta che in tutte le regioni (anche quelle che non hanno esercitato l'opzione di cui al comma 11) sono attivi sistemi informativi paralleli tanto per le prestazioni specialistiche quanto per le prescrizioni farmaceutiche.

Se torniamo ora a concentrarci sulla produzione di misure di interesse epidemiologico, emerge abbastanza chiaramente come ciascuna azienda USL debba fare riferimento, a tal fine, alla popolazione anagraficamente residente nel territorio di propria competenza.

E' pertanto necessario che le ASL vengano messe in condizione di recuperare le informazioni sulle prestazioni erogate ai propri residenti in "mobilità passiva" e cioè da strutture non appartenenti alla ASL stessa. Il peso relativo di tali informazioni varia a seconda delle prestazioni considerate ed assume valori particolarmente alti in presenza di grandi strutture erogatrici "aziendalizzate" che trasmettono i dati direttamente al livello regionale. I dati della mobilità intraregionale vengono di norma forniti alle ASL, con cadenza al massimo annuale, dalle strutture che a livello regionale coordinano il sistema informativo. I dati della mobilità interregionale sono invece oggetto di uno specifico flusso tra regioni, che si avvia entro il mese di luglio dell'anno successivo e si definisce entro il 31 dicembre dello stesso anno.

Per quanto riguarda i contenuti informativi dei sistemi di rilevazione corrente delle "prestazioni", occorre rilevare che le attività del progetto ministeriale "Mattoni del SSN" hanno prodotto e stanno continuando a produrre risultati importanti nell'ottica dell'uso epidemiologico dei dati amministrativi (vedi, per riferimento, <http://www.nsis.salute.gov.it/nsis/nsis.jsp>) I tracciati record e gli schemi di codifica delle rilevazioni più significative sono stati sottoposti ad un utile lavoro di revisione ed omogeneizzazione. Diverse Regioni stanno sperimentando l'introduzione del campo "diagnosi/sospetto diagnostico" nel tracciato record della specialistica ambulatoriale, mentre in altre è stato attivato nelle SDO il codice PAA (presente all'accettazione) che permette di distinguere le patologie concomitanti dalle complicanze. Sono stati inoltre istituiti e regolamentati a livello nazionale il Sistema Informativo per il monitoraggio delle prestazioni in emergenza/urgenza, che dovrebbe tra l'altro documentare le condizioni rilevate e le procedure eseguite in occasione degli accessi in Pronto Soccorso non seguiti da ricovero, il flusso informativo sulle prestazioni farmaceutiche effettuate in distribuzione diretta o per conto, il Sistema Informativo per il monitoraggio delle prestazioni erogate nell'ambito dell'assistenza residenziale e semiresidenziale e il Sistema Informativo per il monitoraggio dell'assistenza domiciliare.

La necessità di riferire le misure di interesse epidemiologico alla popolazione residente impone di estendere la riflessione alla disponibilità, presso le aziende USL, di archivi contenenti dati aggiornati ed esaustivi su tale popolazione.

In teoria, non dovrebbe essere troppo difficile prevedere che ogni USL disponga di informazioni aggiornate e storicizzate (e cioè comprensive delle posizioni cessate, con data e motivo della cessazione, oltre che di quelle attive) sulle persone residenti nei comuni/municipi che ad essa afferiscono.

In pratica, sono assai rare le situazioni in cui sistemi di questo tipo sono stati attivati e funzionano correntemente. Per le USL è infatti decisamente impegnativo garantirsi la segnalazione in continuo, da parte degli uffici comunali, della totalità delle variazioni di interesse: iscrizioni (per nascita o immigrazione), cancellazioni (per morte o emigrazione) e cambi di domicilio (nello stesso comune).

In molte Regioni l'anagrafica degli assistiti rappresenta in realtà solo il supporto alle procedure di scelta e revoca del MMG/PLS. Conseguentemente, anche là dove sono state attivate procedure di acquisizione periodica e riscontro delle variazioni registrate presso l'Agenzia delle Entrate, le informazioni dettagliate sulla residenza (indirizzo e sezione di censimento) e il complesso delle informazioni relative alle persone che non hanno effettuato la scelta del MMG presentano significativi limiti di validità.

Inoltre, l'uso prevalente a cui i dati sono destinati fa sì che non ne venga prevista in modo routinario la "storicizzazione", con le prevedibili conseguenze sulla possibilità di definire il corretto denominatore a cui riferire gli eventi considerati.

D'altra parte, anche là dove i dati anagrafici sono disponibili ed utilizzabili, nella scelta dei denominatori bisogna tenere conto del disallineamento tra risultanze anagrafiche e dati Istat sulla popolazione residente. L'entità di questo disallineamento, che dipende essenzialmente da problemi di copertura delle operazioni di Censimento e dalla mancata cancellazione anagrafica di persone decedute o emigrate, è variabile da comune a comune, ma in alcune situazioni assume una notevole rilevanza (e.g. a Roma, lo scarto tra le 2 popolazioni era di oltre 260000 persone fino al 2005 ed è stato ridotto a circa 120000 a seguito delle verifiche anagrafiche fatte nel 2006, verifiche che hanno portato ad un aumento di oltre il 6% del dato Istat sulla popolazione residente).

L'anagrafe degli assistiti rappresenta nella maggior parte dei casi anche l'archivio di riferimento la gestione informatizzata delle esenzioni dal ticket. Purtroppo, sebbene questa informazione si sia rilevata utile ad integrare i dati ospedalieri e farmaceutici nella stima della prevalenza di malattie croniche, solo in alcune Regioni le aziende USL hanno la possibilità di scaricare l'elenco completo degli esenti per codice.

Concludiamo la rassegna sulle fonti informative a cui fare riferimento per la produzione di misure epidemiologiche a livello di USL soffermandoci sui dati di mortalità.

Per quanto riguarda l'informazione sull'evento morte (che si riduce sostanzialmente alla data di decesso) valgono le già riportate riflessioni sulla necessità di anagrafi aziendali degli assistiti aggiornate e storicizzate. Il recupero diretto di questo dato dalle anagrafi comunali risolverebbe anche il problema di possibili errori nei dati anagrafici al momento della compilazione della scheda Istat. Per quanto riguarda invece il recupero dell'informazione codificata sulla causa principale di morte, i modelli organizzativi e l'interazione con l'Istat possono essere i più diversi, ma l'obiettivo da perseguire è la disponibilità, per la USL, di un archivio contenente un record (comprensivo del codice ICD relativo alla causa principale di morte) per ogni decesso di cittadino residente nella ASL. Si presuppone

pertanto un sistema informativo, aziendale o regionale, che permetta il recupero dell'informazione sui decessi "in mobilità". Vanno infine segnalati la forte disomogeneità tra le regioni rispetto al passaggio al nuovo sistema di codifica delle cause di morte (dalla ICD-9 alla ICD-10) e i problemi posti da tale passaggio (non esiste una tabella di conversione "biunivoca" tra i due sistemi).

### **Gestione delle informazioni**

Abbiamo già detto che, sebbene l'uso "epidemiologico" dei dati amministrativi sia, in teoria, quello centrale rispetto alle funzioni attribuite alle aziende USL dall'ordinamento corrente, lo stimolo principale alla raccolta di tali dati è dato da esigenze di tipo gestionale, quando non meramente contabile.

Pertanto, se da una parte le Unità Operative di Epidemiologia (comunque esse vengano etichettate nell'organigramma aziendale) devono pretendere di partecipare attivamente, evidenziando le proprie esigenze, alla fase di progettazione dei sistemi di archiviazione ed interrogazione dei database, dall'altra non si può ignorare che dagli archivi di cui stiamo ipotizzando la costruzione "pescherebbero" anche, per le loro esigenze, le componenti "amministrativo-contabile" e "gestionale" dell'azienda.

Al proposito, sono possibili diverse soluzioni organizzative e tecnologiche per soddisfare le necessità informative delle diverse articolazioni aziendali:

- Un unico "*data repository*" che viene alimentato direttamente dalle procedure di supporto alla produzione e a cui attingono tutti, ognuno per ricavare i propri indicatori
- Una struttura a 2 livelli, in cui le procedure di supporto alla produzione generano direttamente gli indicatori gestionali "tematici", mentre il caricamento dei microdati nel "*data repository* riconciliato" avviene in un momento temporalmente e logicamente distinto, ha come oggetto primario la popolazione dei residenti nel territorio aziendale (indipendentemente dalla struttura che ha erogato la prestazione), è subordinato ad un più stretto controllo di qualità ed è principalmente orientato alla produzione di misure di tipo epidemiologico (ma può servire anche alla produzione di indicatori gestionali che richiedano l'integrazione di dati di diversa provenienza).

La seconda soluzione sembra preferibile, perché si pone l'obiettivo di soddisfare esigenze diverse in termini di "estensione", "rigore" e "tempestività". Si sconsigliano invece soluzioni in cui misure epidemiologiche vengano calcolate routinariamente attraverso procedure automatiche e incluse nei "cruscotti" messi a disposizione della direzione aziendale.

Inoltre, soprattutto in caso di esternalizzazione dei servizi informatici, è importante che i dati siano organizzati in modo che l'area "epidemiologica" possa accedervi direttamente per effettuare elaborazioni non predefinite (e.g. utilizzando SAS o Stata).

Le attività di integrazione degli archivi informatizzati devono fare i conti con i vincoli imposti dalla normativa sulla riservatezza dei dati personali sensibili. Tali vincoli sono meno stringenti a livello di azienda USL di quanto non lo siano, per esempio, a livello regionale. E' del resto ovvio che tanto le stime di occorrenza quanto gli altri indicatori di tipo epidemiologico che le USL ritengano opportuno calcolare richiedono il collegamento tra le diverse informazioni (presenti all'interno dello stesso flusso o di flussi diversi) relative alla stessa persona. Nel caso della produzione di misure di processo e di esito è necessaria una vera e propria ricostruzione longitudinale, in capo ai singoli individui, dei percorsi assistenziali e/o delle storie di malattia. E' quindi indispensabile che i microdati utilizzati per tali

elaborazioni contengano identificativi individuali. E' però opportuno che nella costruzione del *data repository* gli identificativi "parlanti" (nome, cognome, sesso, luogo e data di nascita, CF) vengano sostituiti da un codice individuale non parlante comune, per lo stesso soggetto, a tutti gli archivi sottoposti ad integrazione e che l'accesso alla tabella di attribuzione di tale codice sia limitato alle operazioni strettamente necessarie e alle persone incaricate di svolgere le operazioni stesse.

## RILEVAZIONE SULLE CARATTERISTICHE DEI SISTEMI INFORMATIVI SANITARI NELLE REGIONI ITALIANE

La rilevazione si è focalizzata sulla possibilità, per le aziende USL, di avere routinariamente a disposizione informazioni analitiche e nominative relative ai propri residenti, dalle quali ricavare misure di interesse epidemiologico.

Il questionario, che era articolato in 7 aree di interesse:

1. anagrafe dei residenti/assistiti
2. esenzioni dal ticket
3. ricoveri ospedalieri
4. specialistica ambulatoriale
5. farmaceutica territoriale
6. distribuzione diretta dei farmaci
7. mortalità per causa

E' stato inviato dal CCM a 14 tra Regioni e Province Autonome, selezionate in base alla "prossimità" con almeno uno dei componenti del gruppo di esperti, che avrebbe quindi potuto favorirne la compilazione.

Otto Regioni e una P.A. hanno riconsegnato il questionario compilato:

- Valle d'Aosta
- Piemonte
- Lombardia
- P.A. Trento
- Emilia-Romagna
- Toscana
- Lazio
- Basilicata
- Sicilia

Questo ha consentito di delineare un quadro abbastanza articolato, seppur non esaustivo, delle diverse modalità di funzionamento dei sistemi di raccolta, validazione e "restituzione" alle aziende dei dati sanitari correnti.

Le domande incluse nel questionario e i risultati analitici della rilevazione vengono riportati al termine del documento (**Allegato C**). Nel seguito vengono messi in evidenza, per le diverse aree tematiche, i principali elementi emersi dall'indagine.

### Anagrafe dei residenti/assistiti

In 8 regioni su 9 le ASL hanno routinariamente a disposizione i dati anagrafici dei propri residenti. Tuttavia, solo in 6 casi questa informazione viene storicizzata e solo in 4 casi il dato sulle posizioni cessate è corredato dal motivo della cessazione. Questa

disomogeneità è probabilmente legata alla grande varietà, tra regione e regione, dei flussi informativi utilizzati per segnalare al servizio sanitario le variazioni anagrafiche.

#### Esenzione dal ticket

Sebbene 8 regioni su 9 dispongano di archivi informatizzati contenenti questa informazione, tali archivi vengono messi routinariamente a disposizione delle aziende USL solo in 3 casi.

#### Ricoveri ospedalieri per acuti e post-acuti

Questo flusso informativo, normato a livello nazionale, è piuttosto omogeneo. Tuttavia, solo in 7 regioni le USL hanno accesso routinario ai dati sui ricoveri dei propri residenti (mobilità intraregionale) e solo in 6 l'accesso è esteso alla mobilità interregionale. Neanche il riscontro dei dati dei pazienti con quelli contenuti nell'anagrafe assistiti avviene in modo generalizzato.

#### Specialistica ambulatoriale

In tutte le Regioni è attivo un sistema informativo "proprio"; in 6 casi è questo sistema informativo a "passare" i dati al sistema TS (opzione di cui al comma 11 dell'art.50), negli altri 3 i due sistemi operano parallelamente e indipendentemente.

E' assai raro che le USL scarichino i dati di propria competenza dal sistema TS (anche perché tali dati, come già chiarito, non si estendono a tutte le prestazioni erogate ai residenti), così come è piuttosto infrequente che i dati anagrafici dei pazienti siano riscontrati sull'anagrafe assistiti. La mobilità passiva interregionale viene resa disponibile alle ASL in 5 regioni su 9.

#### Farmaceutica territoriale

La situazione è analoga a quella della specialistica ambulatoriale. In questo caso, tuttavia, solo una Regione si è avvalsa del comma 11 mentre nelle altre 8 il sistema TS opera parallelamente al sistema regionale di acquisizione ottica delle ricette. La mobilità passiva interregionale viene resa disponibile alle ASL in 2 regioni su 9.

#### Distribuzione diretta dei farmaci

Il flusso informativo è attivato in tutte le regioni con modalità sostanzialmente omogenee. Nella maggior parte dei casi le USL possono disporre dei dati analitici relativi ai residenti anche se la correttezza dei dati anagrafici viene verificata solo in 6 casi su 8 e la mobilità interregionale resa disponibile solo in 3 Regioni.

#### Mortalità con indicazione della causa

In 8 delle 9 Regioni considerate vengono tenuti Registri informatizzati delle Cause di Morte e in 7 l'informazione relativa ai residenti è resa routinariamente disponibile alle USL. Ci sono tuttavia importanti disomogeneità, che riguardano le procedure di validazione dei dati anagrafici, il sistema di codifica adottato (5 ICD-9 e 3 ICD-10) e il livello a cui la codifica stessa viene eseguita.

L'immagine complessiva che si può trarre dalla rilevazione effettuata è quella di tanti modelli di sistema informativo sanitario quante sono le Regioni e le P.A. considerate. Salvo rare eccezioni, questi modelli sembrano derivare più dall'accumularsi di esigenze informative di tipo diverso che dalla formulazione di un disegno unitario. In ogni caso le finalità di tipo amministrativo/contabile (supporto ai sistemi di finanziamento) e quelle legate al monitoraggio della spesa sanitaria sembrano prevalere rispetto a quelle di osservazione epidemiologica e di valutazione degli effetti degli interventi sanitari sulla salute della popolazione. In questo quadro, le esigenze informative individuate dalle aziende USL come presupposto per un corretto svolgimento della propria funzione di tutela non sembrano rappresentare una priorità per i livelli regionali di progettazione ed organizzazione dei sistemi informativi correnti. Anche la semplice disponibilità di una anagrafe completa, aggiornata e storicizzata dei residenti, che rappresenterebbe il punto di partenza per qualsiasi tentativo di approfondimento epidemiologico è, nella maggior parte dei casi, tutt'altro che garantita.

## ALLEGATO C

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

### RILEVAZIONE SULLE CARATTERISTICHE DEI SISTEMI INFORMATIVI SANITARI NELLE REGIONI ITALIANE

#### Note per la compilazione

Un archivio è considerato “routinariamente disponibile” per la ASL quando, senza necessità di richieste specifiche, la ASL può disporre di un file che può essere caricato senza eccessive complicazioni su un DBMS (Access, Oracle, etc.) ed ivi interrogato con query SQL. Il file deve pertanto essere corredato di un tracciato record contenente nome, lunghezza, tipo e schema di codifica delle variabili registrate. Per essere considerati “disponibili” gli archivi devono pertanto essere “scaricabili” e non solamente “interrogabili”. Il questionario fa riferimento ad informazioni individuali analitiche (cioè relative a singoli individui e a singoli contatti con i servizi sanitari) e nominative (cioè corredate di uno o più identificativi individuali che permettono il collegamento tra i diversi flussi e quello a fonti informative esterne).

#### 1. Anagrafe assistiti

1.1	La Regione cura la gestione e l'aggiornamento dell'anagrafe regionale dei residenti/assistiti?	<b>1 NO</b>	<b>8 SI</b>
1.2	Se SI alla domanda precedente:		
	1.2.1 A parte le segnalazioni registrate direttamente dagli sportelli per la scelta e revoca del MMG e PLS, quale è il flusso informativo delle variazioni anagrafiche relative ai residenti nei comuni della regione (nascite, immigrazioni, decessi, emigrazioni cambiamenti di residenza tra comuni della regione)? <ul style="list-style-type: none"> <li>• dai comuni direttamente al livello regionale <span style="float: right;">① 2</span></li> <li>• dai comuni alle ASL e da queste al livello regionale <span style="float: right;">② 5</span></li> <li>• dai Comuni all'Agenzia delle Entrate e da questa al livello regionale <span style="float: right;">③ 5</span></li> <li>• dai Comuni all'Agenzia delle Entrate, da questa alle ASL e da queste al livello regionale <span style="float: right;">④ 1</span></li> </ul>		
	1.2.2 Quale è, se diverso, il flusso informativo dei cambiamenti di residenza all'interno dello stesso comune? solo su segnalazione diretta ① 1      ①      ②      ③      ④		
	1.2.3 Alle ASL vengono resi routinariamente disponibili (scaricabili, non solo interrogabili) i dati analitici dell'anagrafe assistiti?	<b>0 NO</b>	<b>8 SI</b>
1.2.4	Se SI alla domanda precedente:		
	1.2.4.1 Alle ASL vengono resi routinariamente disponibili i dati relativi a: <ul style="list-style-type: none"> <li>• le posizioni attive (fotografia) <span style="float: right;">① 2</span></li> <li>• le posizioni attive e le posizioni cessate (archivio storico) <span style="float: right;">② 6</span></li> </ul>		

	1.2.4.2 I record delle posizioni attive contengono la data di decorrenza della posizione?	1 NO	7 SI
	1.2.4.3 I record delle posizioni attive contengono il motivo di decorrenza della posizione (nascita, iscrizione, cambio di MMG, etc.)?	5 NO	3 SI
	1.2.4.4 I record delle posizioni cessate (se disponibili) contengono la data di cessazione della posizione?	0 NO	6 SI
	1.2.4.5 I record delle posizioni cessate (se disponibili) contengono il motivo di cessazione della posizione (morte, emigrazione, cambio di MMG, etc.)?	2 NO	4 SI
	1.2.4.6 Le informazioni rese disponibili alle ASL sono relative a: <ul style="list-style-type: none"> <li>• tutti i residenti della regione + gli assistiti non residenti ① 2 N.A. 2 (*)</li> <li>• tutti i residenti nella ASL + gli assistiti non residenti ② 4</li> </ul>		

(\*) Per la Valle d'Aosta e la P.A. di Trento la ASL coincide con la Regione/P.A.

## 2. Esenzioni ticket

2.1	E' attivo un sistema di gestione informatizzata delle esenzioni dal ticket a livello regionale?	1 NO	8 SI
2.2	Se SI alla domanda precedente:		
	2.2.1 L'archivio è integrato con l'anagrafe regionale dei residenti/assistiti?	1 NO	6 SI 1 MI
	2.2.2 L'informazione analitica e nominativa sui residenti esenti per codice viene resa routinariamente disponibile (scaricabile, non solo interrogabile) alle ASL?	4 NO	3 SI 1 MI
	2.2.3 Se si, con quale cadenza?		
	2.2.4 L'archivio delle esenzioni attive contiene la data di inizio esenzione?	0 NO	7 SI 1 MI
	2.2.5 I dati analitici e nominativi sulle esenzioni cessate con data e motivo di cessazione vengono resi disponibili alle ASL?	4 NO	3 SI 1 MI



#### 4. Flusso informativo sulle prestazioni specialistiche ambulatoriali

4.1	E' attivato il flusso informativo "ex art.50 L.326/2003 " verso il MEF/SOGEI?	<b>0 NO</b>	<b>9 SI</b>
4.2	Se SI alla domanda precedente:		
	4.2.1 Qual è la copertura del flusso in termini di strutture rispondenti (I trimestre 2009)?	Dal 70% al 100%	
	4.2.2 Qual è la copertura del flusso in termini di prestazioni registrate (I trimestre 2009) ?	Dal 70% al 100%	
	4.2.3 Le ASL recuperano routinariamente da MEF/SOGEI (utilizzando l'applicativo XFB) i dati analitici relativi alle: <ul style="list-style-type: none"> <li>• prestazioni erogate dalle proprie strutture <b>① 1 SI 8 NO</b></li> <li>• prestazioni erogate dalle strutture accreditate ubicate nel proprio territorio <b>② 8 NO</b></li> <li>• prestazioni ovunque erogate ai propri assistiti/residenti <b>③ 9 NO</b></li> </ul>		
	4.2.4 La regione si è avvalsa dell'opzione di cui al comma 11 del citato art.50?	<b>3 NO</b>	<b>6 SI</b>
4.3	Indipendentemente dal flusso ex art.50, è attivo un sistema informativo regionale sull'assistenza specialistica ambulatoriale che rileva dati analitici sulle singole prestazioni erogate?	<b>0 NO</b>	<b>9 SI</b>
4.4	Se SI alla domanda precedente:		
	4.4.1 Con che cadenza i dati vengono inviati dai soggetti erogatori alla regione?	<b>2 gior, 5 mens, 1 trim., 1 a richiesta</b>	
	4.4.2 I dati vengono inviati alla regione nella modalità "file disgiunti" (file dei dati anagrafici e file dei dati sanitari collegati tramite chiave univoca)?	<b>2 NO</b>	<b>7 SI</b>
	4.4.3 La presenza nell'anagrafe assistiti dei pazienti indicati come residenti: <ul style="list-style-type: none"> <li>• non viene verificata <b>① 3 + 2 MI</b></li> <li>• viene verificata da parte del soggetto erogatore <b>② 2</b></li> <li>• viene verificata a livello regionale <b>③ 2</b></li> </ul>		
	4.4.4 Se 2 o 3 alla domanda precedente, viene verificata, ed eventualmente corretta, l'informazione su ASL/Comune/Municipio di residenza?	<b>3 NO</b>	<b>1 SI</b>
	4.4.5 E' possibile, nei files inviati alla regione, distinguere i pazienti identificati attraverso la lettura ottica della carta sanitaria?	<b>8 NO</b>	<b>0 SI 1 MI</b>

	4.4.6 In che data l'archivio regionale (non inclusivo della mobilità passiva interregionale) relativo all'anno precedente viene considerato definitivo e reso disponibile alle ASL?		
		<b>1 NO, da 2 mesi a 1 anno</b>	
	4.4.7 Quali parti di tale archivio vengono routinariamente rese disponibili alle ASL?		
	<ul style="list-style-type: none"> <li>• prestazioni erogate dalle strutture a gestione diretta ① 7</li> <li>• prestazioni erogate dalle strutture accreditate localizzate sul territorio della ASL ② 8</li> <li>• prestazioni prescritte da medici della ASL e da MMG e PLS convenzionati con la ASL ③ 3</li> <li>• prestazioni erogate nei confronti dei residenti della ASL ④ 7</li> </ul>		
4.5	Il livello regionale rende disponibili alle ASL i dati analitici e nominativi sulle prestazioni erogate fuori regione a persone residenti (mobilità passiva interregionale)?	<b>4 NO</b>	<b>5 SI</b>
4.6	Se SI alla domanda precedente:		
	4.6.1 In che data?	<b>Tra 7 e 10 mesi</b>	
	4.6.2 Nello stesso formato dei punti 4.4.6 e 4.4.7?	<b>2 NO</b>	<b>3 SI</b>

## 5. Flusso informativo sulle prescrizioni farmaceutiche territoriali

5.1	E' attivato il flusso informativo "ex art.50 L.326/2003 " verso il MEF/SOGEI?	<b>0 NO</b>	<b>9 SI</b>
5.2	Se SI alla domanda precedente:		
	5.2.1 Qual è la copertura del flusso in termini di farmacie rispondenti (I trimestre 2009)?	<b>6 - 90/100%</b> <b>3 - MI</b>	
	5.2.2 Qual è la copertura del flusso in termini di prescrizioni registrate (I trimestre 2009)?	<b>6 - 90/100%</b> <b>3 - MI</b>	
	5.2.3 Le ASL recuperano routinariamente da MEF/SOGEI (utilizzando l'applicativo XFB) i dati analitici relativi alle:		
	• prescrizioni spese presso farmacie ubicate nella ASL	<b>① 4 NO</b>	<b>5 SI</b>
	• prescrizioni, ovunque spese, emesse da medici della ASL e da MMG e PLS convenzionati con la ASL	<b>② 8 NO</b>	<b>1 SI</b>
	• prescrizioni ovunque spese dai propri assistiti/residenti	<b>③ 8 NO</b>	<b>1 SI</b>
	5.2.4 La regione si è avvalsa dell'opzione di cui al comma 11 del citato art.50?	<b>8 NO</b>	<b>1 SI</b>
5.3	Indipendentemente dal flusso ex art.50, è attivo un sistema informativo regionale sull'assistenza farmaceutica che rileva dati analitici e nominativi sulle singole prescrizioni spese?	<b>0 NO</b>	<b>9 SI</b>
5.4	Se SI alla domanda precedente:		
	5.4.5 Con che cadenza, e a distanza di quanto tempo, i dati vengono resi disponibili alla regione?	<b>1 g., 6 m. 2 t.</b> <b>0-8 mm.</b>	
	5.4.6 La presenza nell'anagrafe assistiti dei pazienti indicati come residenti:		
	• non viene verificata	<b>① 2</b>	
	• viene verificata da parte della farmacia	<b>② 1 + 2 MI</b>	
	• viene verificata dalla ASL	<b>③ 1</b>	
	• viene verificata al livello regionale	<b>④ 3</b>	
	5.4.7 Se 2, 3 o 4 alla domanda precedente, viene verificata, ed eventualmente corretta, l'informazione su ASL/Comune/Municipio di residenza?	<b>5 NO</b>	<b>0 SI</b>
	5.4.8 E' possibile, nei files inviati alla regione, distinguere i pazienti identificati attraverso la lettura ottica della carta sanitaria?	<b>8 NO</b>	<b>0 SI</b> <b>1 MI</b>
	5.4.9 Quali parti dell'archivio regionale vengono routinariamente rese disponibili alle ASL?		





## 7. Flusso informativo sulla mortalità per causa

7.1	E' attivo un sistema informativo regionale finalizzato alla realizzazione del Registro Nominativo delle Cause di Morte (ReNCaM)?	1 NO	8 SI
7.2	Se SI alla domanda precedente:		
	7.2.1. I dati anagrafici dei deceduti registrati nel ReNCaM: <ul style="list-style-type: none"> <li>• provengono dalle anagrafi comunali ① 1 (Roma)</li> <li>• provengono dalla scheda di morte Istat, ma sono riscontrati con quelli delle anagrafi comunali ② 2</li> <li>• provengono dalla scheda di morte Istat e non sono riscontrati con quelli delle anagrafi comunali ③ 6</li> </ul>		
	7.2.2 Quale classificazione viene usata per codificare la causa principale di morte? <ul style="list-style-type: none"> <li>• ICD-9 ① 5</li> <li>• ICD-10 ② 3</li> </ul>		
	7.2.3 Dove avviene la codifica della causa di morte? <ul style="list-style-type: none"> <li>• presso le ASL ① 6</li> <li>• a livello regionale ② 1 + Roma</li> <li>• presso l'Istat ③ 1</li> </ul>		
	7.2.4 Sono disponibili per le ASL i ReNCaM inclusivi della informazione relativa a tutti i residenti deceduti (morti in comuni della ASL, in altri comuni della regione, fuori regione)?	1 NO	7 SI
	7.2.5 Se SI, in che data?	1-2 anni	

## **RELAZIONE GRUPPO DI LAVORO AREA 2: UTILIZZO DI BASI DI DATI INTEGRATI<sup>1</sup>**

Il gruppo di lavoro è partito dalla esperienza di definizione di algoritmi per la identificazione di casi incidenti o prevalenti di specifiche patologie, maturata nel corso del precedente progetto e riassunta nella monografia di Epidemiologia e Prevenzione (Epidemiologia & Prevenzione 2008; 32 Supplemento maggio-giugno, ed ha svolto la propria attività secondo tre linee di lavoro:

1. Approfondimento di alcuni algoritmi, già presentati nella monografia.
2. Definizione di algoritmi, sempre basati sui dati disponibili da fonti sanitarie correnti, per la identificazione di casi incidenti o prevalenti di malattie che non erano state prese in precedenza in esame.
3. Utilizzo dei casi di patologia identificati attraverso gli algoritmi presentati nella monografia di Epidemiologia e Prevenzione, per la valutazione dei percorsi assistenziali dei pazienti affetti dalla malattia.

Il primo punto ha visto la proposta di nuovi algoritmi per la stima della incidenza delle Sindromi Coronariche Acute (che rappresenta una estensione dell'algoritmo di stima della incidenza dell'infarto miocardico acuto) e la revisione di quello per la stima della BPCO (che cerca di superare i limiti evidenziati per la definizione di BPCO nel precedente lavoro, quali la sottostima dei casi BPCO e la selezione di casi BPCO medio-gravi).

Il secondo punto ha visto la definizione di algoritmi per la stima di ipo- ed iper-tiroidismo.

Il terzo punto ha utilizzato i casi di diabete, identificati secondo l'algoritmo del precedente lavoro per valutare il follow-up ambulatoriale e gli esiti a distanza dei pazienti, sempre con dati sanitari correnti.

---

<sup>1</sup> Il gruppo di lavoro è stato coordinato da Alessandro Barchielli e da Roberto Gnani

## 1. Estensione dell'algoritmo di stima dell'infarto miocardico acuto all'insieme delle Sindromi Coronariche Acute<sup>2</sup>

L'algoritmo utilizzato per la stima della incidenza dell'infarto miocardico acuto (IMA) nel precedente progetto<sup>3</sup>, basato sugli archivi delle schede di dimissione ospedaliera (SDO) e sui registri locali delle cause di morte (CM) è riportato nella tabella seguente.

Fonte dei dati e criteri di definizione di caso incidente di IMA <sup>3</sup>		
Fonte	Criteri selezione casistica	Criteri definizione incidenza
<b>Mortalità</b>	Decesso per infarto miocardico acuto in diagnosi di morte principale	Assenza di altri ricoveri con diagnosi di dimissione principale o secondaria con i codici ICD-9-CM 410 o 412 (infarto miocardico pregresso), nei 5 anni precedenti alla data di ammissione o di morte
<b>Schede di dimissione ospedaliera</b>	Ricovero ordinario per infarto miocardico acuto (ICD-9-CM: 410) in diagnosi di dimissione principale o secondaria, se associata ad alcuni specifici codici in diagnosi principale <sup>1</sup>	

(1) Codici in diagnosi di dimissione principale: 427.1, 427.41, 427.42, 427.5, 428.1, 429.5, 429.6, 429.71, 429.79, 429.81, 518.4, 780.2, 785.51, 414.10, 423.0

Siamo partiti da tale esperienza per cercare di mettere a punto e valutare un algoritmo di stima di tutte le forme Sindromi Coronariche Acute (SCA) partendo da due considerazioni:

1. L'importanza sul piano clinico della angina instabile (AI), e le problematiche di diagnosi differenziale tra AI ed alcune forme di IMA, con estensione limitata del danno ischemico. Infatti la introduzione nella pratica clinica corrente, avvenuta nei primi anni 2000, del test della troponina che è in grado di evidenziare fenomeni molto limitati di necrosi miocardica e la definizione di nuovi criteri di diagnosi di

<sup>2</sup> A cura di: Alessandro Barchielli, Roberto Gnani e Roberta Picariello

<sup>3</sup> Barchielli A, Balzi D, Bruni A et al. Stima dell'incidenza dell'infarto miocardico acuto basata sui dati sanitari correnti mediante l'uso di un algoritmo comune in differenti aree italiane. *Epidemiol Prev.* 2008;32(3 Suppl):30-7.

IMA che valorizzano il ruolo delle modificazioni enzimatiche, hanno portato a modifiche nella differenziazione tra AI ed IMA che possono essere avvenute in tempi diversi nelle diverse tipologie di strutture sanitarie e nelle varie aree del paese.

2. La difficoltà della diagnosi di IMA in caso di decesso extraospedaliero ed in mancanza di un riscontro autoptico (che viene effettuato molto raramente), rende poco affidabile la diagnosi di IMA in caso, appunto, di morte fuori dall'ospedale. Queste difficoltà avevano già portato, il progetto MONICA ad introdurre il concetto di Morte Coronarica.

## Nuovo algoritmo per la definizione di Sindrome Coronarica Acuta

La tabella seguente riporta il nuovo algoritmo con la definizione dei casi incidenti di SCA, che:

1. Introduce i ricoveri con diagnosi principale di dimissione per il codice ICD-9 CM 411, per identificare i casi di AI.
2. Utilizza un criterio di priorità clinica nella definizione della incidenza di IMA e di AI, che tenendo conto della maggior gravità clinica dell'IMA, attribuisce al ricovero per IMA una priorità rispetto a quello per AI.
3. Estende la mortalità all'insieme dei decessi per 410-414, definite come morti coronariche.

Il nuovo algoritmo è stato applicato a Firenze ed a Torino.

Fonte dei dati e criteri di definizione di caso incidente di SCA			
Fonte	Tipo evento	Criteri selezione casistica	Criteri definizione incidenza
SDO	<b>Infarto Miocardico acuto</b>	Ricovero ordinario <u>con 410 *</u> in diagnosi di dimissione principale o secondaria, se associata ad alcuni specifici codici in diagnosi principale <sup>1</sup>	Assenza di altri ricoveri con diagnosi di dimissione principale o secondaria per 410* o per 412*, nei 5 anni precedenti alla data di ammissione
	<b>Angina Instabile</b>	Ricovero ordinario <u>con 411</u> in diagnosi di dimissione principale o secondaria, se associata ad alcuni specifici codici in diagnosi principale <sup>1</sup>	Assenza di altri ricoveri con diagnosi di dimissione principale o secondaria per 410*, 411 o 412*, nei 5 anni precedenti alla data di ammissione
<b>Mortalità</b>	<b>Morte coronarica</b>	Decesso <u>con 410*-414*</u> in diagnosi di morte principale	Assenza altri ricoveri con diagnosi di dimissione principale o secondaria per 410*, 411 o 412*, nei 5 anni precedenti alla data di ammissione o di morte

(1) Codici in diagnosi di dimissione principale: 427.1, 427.41, 427.42, 427.5, 428.1, 429.5, 429.6, 429.71, 429.79, 429.81, 518.4, 780.2, 785.51, 414.10, 423.0

## Risultati

Le tabelle 1-3 riportano i risultati dell'applicazione dei nuovi algoritmi di definizione dei casi incidenti di SCA effettuate a Firenze e Torino.

La tabella 1 riporta il confronto dei tassi standardizzati per età (standard: popolazione Italia 2001, totale) per il totale di maschi e femmine, osservati nelle 2 aree. Il confronto tra Firenze e Torino può essere fatto per gli anni comuni (2002-2004). I dati per le diverse forme di SCA sono abbastanza stabili nei tre anni. Nel confronto tra i tassi delle 2 aree si osservano valori dello stesso ordine di grandezza per l'IMA ospedalizzato e per le Morti Coronariche. Viceversa, Firenze presenta un valore più elevato di circa il 30% dei tassi di incidenza di AI, che si riflette in un valore più elevato dei tassi del totale delle SCA. Tale differenza non si presenta di facile interpretazione. Da un punto di vista generale possono essere avanzate varie ipotesi.

1. La differenza di incidenza è reale. Da un punto di vista generale è plausibile osservare una diversa incidenza della patologia coronarica tra diverse aree, ma è più difficile spiegare come la differenza possa concentrarsi solo in un specifica forma ischemica.
2. Le differenze osservate sono il frutto di differenze nei criteri di diagnosi clinica e/o in quelli di codifica della diagnosi di dimissione. Questa in particolare potrebbe riguardare la classificazione tra angina instabile ed angina stabile (cod. 413; AS)

Per approfondire questa ipotesi sono state fatte due verifiche, confrontando:

- la mortalità ospedaliera di IMA e AI ospedalizzate nelle due aree (tabella 2);
- i tassi di ospedalizzazione per AI ed AS (tabella 3).

La mortalità ospedaliera per i codici 410+411 è in tutti gli anni più elevata a Torino, anche se l'età media dei pazienti è più giovane di circa 2 anni. La differenza è particolarmente concentrata nel codice 411: a Torino è circa doppia di quella di Firenze, sempre con pazienti mediamente più anziani rispetto a Firenze.

Questa differenza potrebbe essere coerente con una maggior tendenza a classificare a Firenze con i codice 411 (AI) casi che a Torino vengono classificati con il codice 413 (AS). Il diverso comportamento porterebbe ad una incidenza più elevata dell'AI a Firenze e ad una sua miglior prognosi, in quanto verrebbero incluse forme meno gravi. Si tratterebbe quindi di una problematica analoga al "fenomeno di Will Rogers" osservato a proposito del rapporto tra modifiche nella stadiazione dei tumori dovuti a miglioramenti diagnostici che portano alla identificazione di metastasi, in precedenza non diagnosticate, e miglioramenti nella prognosi dei tumori. Il confronto della distribuzione percentuale dei ricovero per 411 e 413 potrebbe, almeno in parte, confermare tale ipotesi. Infatti la proporzione di ospedalizzazioni per 411 è più elevata a Firenze (tabella 3).

Tabella 1. Tassi standardizzati di incidenza delle varie forme di SCA a Torino e Firenze.

		2002	2003	2004	2005	2006
<b>Totale SCA</b>	<b>Torino</b>	268,9	283,5	267,9		
	<b>Firenze</b>	308,9	307,0	296,6	302,3	287,7
<b>IMA ospedalizzato</b>	<b>Torino</b>	134,2	135,7	140,0		
	<b>Firenze</b>	144,7	144,9	137,3	146,7	140,0
<b>Angina Instabile ospedalizzata</b>	<b>Torino</b>	71,9	77,2	69,9		
	<b>Firenze</b>	102,4	99,5	105,1	103,3	92,3
<b>Morti Coronariche</b>	<b>Torino</b>	64,7	72,3	59,5		
	<b>Firenze</b>	61,8	62,5	54,2	52,4	46,4

Tabella 2. Mortalità ospedaliera grezza per i casi di Infarto Miocardico Acuto (IMA) e di Angina Instabile (AI) osservata a Firenze e Torino

DIAGNOSI PRINCIPALE	Età media	FIRENZE				
		2002	2003	2004	2005	2006
IMA (410)	72,4	9,39	8,94	8,51	7,60	7,11
AI (411)	70,5	0,86	0,65	0,34	0,55	0,61
<b>IMA+ AI (410+411)</b>	69,3	<b>5,04</b>	<b>4,75</b>	<b>4,22</b>	<b>4,03</b>	<b>3,99</b>
DIAGNOSI PRINCIPALE	Età media	TORINO				
		2002	2003	2004	2005	2006
IMA (410)	69,7	11,5	10,4	8,33		
AI (411)	68,1	1,45	1,34	1,86		
<b>IMA+ AI (410+411)</b>	67,4	<b>6,67</b>	<b>6,09</b>	<b>5,38</b>		

Tabella 3. Proporzione di ricoveri Angina Instabile (AI) ed Angina Stabile (AS) osservata a Firenze e Torino

DIAGNOSI PRINCIPALE	FIRENZE				
	2002	2003	2004	2005	2006
	(N=2.623)	(N=2.384)	(N=2.679)	(N=2.603)	(N=2.350)
	%	%	%	%	%
AI (411)	71,10	77,47	76,37	76,22	77,02
AS (413)	28,90	22,53	23,63	23,78	22,98
AI+AS (411+413)	100,0	100,0	100,0	100,0	100,0
DIAGNOSI PRINCIPALE	TORINO				
	2002	2003	2004	2005	2006
	(N=2.437)	(N=2.530)	(N=2.526)		
	%	%	%		
AI (411)	65,2	64,8	61,8		
AS (413)	34,8	35,2	38,2		
AI+AS (411+413)	100,0	100,0	100,0		

### Conclusioni

Prima di trarre indicazioni definitive il nuovo algoritmo dovrebbe essere testato in altre realtà per ampliare i confronti e verificare quanto osservato a Torino. Comunque dai dati presentati emerge, almeno negli anni in esame, una possibile influenza di diversi criteri di diagnosi o di codifica tra le due aree in studio, con possibili misclassificazioni tra AI ed AS. Solo uno studio di validazione delle diagnosi di dimissione effettuata con criteri standardizzati può aiutare a chiarire tale aspetto.

## 2. Revisione dell'algoritmo per la stima di prevalenza della broncopneumopatia cronico ostruttiva (BPCO).<sup>4</sup>

L'algoritmo utilizzato per la stima di prevalenza della BPCO (Epidemiologia e Prevenzione 2008; 32 Supplemento maggio-giugno) si basava sulla integrazione di due archivi elettronici: il registro delle schede di dimissione ospedaliera (SDO) e il registro regionale o comunale delle cause di morte (CM)<sup>4a</sup>.

Riportiamo l'algoritmo completo, nella tabella che segue.

<b>Fonte dei dati e criteri di definizione di caso prevalente di BPCO</b>		
<b>Fonte</b>	<b>Criteri selezione</b>	<b>Criteri definizione caso prevalente</b>
SDO	dimissione con diagnosi principale o secondaria di BPCO, definita mediante codici ICD-9 CM 490 - 492, 494, 496	dimissione con diagnosi di BPCO o decesso per BPCO nell'anno di interesse  oppure  dimissione con diagnosi di BPCO in uno dei 4 anni precedenti se vivi al 1 gennaio dell'anno di interesse
CM	decesso per una causa identificata mediante codici ICD-9 CM 490 - 492, 494, 496	

Punto di forza di questo algoritmo era l'approccio longitudinale, che permetteva di includere nella stima di prevalenza annuale i casi ospedalieri di BPCO osservati negli anni precedenti; questo approccio ha permesso di aumentare la sensibilità delle stime annuali della prevalenza della BPCO di due-tre volte rispetto alla prevalenza stimata con approccio trasversale.

<sup>4</sup> A cura di: Annunziata Faustini, Silvia Cascini, Cristina Canova, Casotto E, Gianfranco De Girolamo, Karim Bonora

<sup>4a</sup> Faustini A, Cascini S, Arcà M, Balzi D, Barchielli A, Canoca C et al. Stima della prevalenza di bronco pneumopatia cronico-ostruttiva basata su dati sanitari correnti, mediante l'uso di un algoritmo comune in differenti aree italiane. Epidemiol Prev. 2008;32(3 Suppl):30-7.

<sup>4b</sup> Anechino C, Rossi E, Fanizza C, De Rosa M, Tognoni G, Romero M. Prevalence of chronic obstructive pulmonary disease and pattern of comorbidities in a general population. Int J COPD 2007; 2: 567-574

La revisione dell'algoritmo trova la suo razionale nella consapevolezza dei limiti che esso presenta, quali la sottostima dei casi BPCO e la selezione di casi BPCO medio-gravi.

La revisione è resa possibile oggi dalla disponibilità di dati ospedalieri per un periodo di osservazione più lungo di 5 anni e dalla disponibilità dei dati delle prestazioni farmaceutiche, relativi all'acquisto di farmaci su prescrizione del medico di base.

Circa l'utilizzo di altre possibili fonti, è stata testata la possibilità di utilizzare l'archivio di esenzione ticket per identificare i casi BPCO dal codice della insufficienza respiratoria. Questa fonte è stata esclusa per il basso contributo che fornisce alla stima di prevalenza, per la possibile misclassificazione dei casi e perché presenta problemi di riproducibilità.

### **OBIETTIVI DELLA REVISIONE**

- 1) stimare l'incremento di sensibilità della prevalenza di BPCO mediante l'uso dei dati della farmaceutica e/o dei dati ospedalieri per periodi di osservazione più lunghi;
- 2) verificare l'ipotesi della selezione di casi medio-gravi di BPCO da fonte ospedaliera, testando la specificità dei casi ospedalieri rispetto ai casi gravi di BPCO come definiti dal consumo di farmaci;
- 3) verificare la possibile misclassificazione dei casi BPCO definiti dal consumo di farmaci, rispetto a casi di asma. Questo terzo obiettivo risponde alla particolare situazione delle patologie respiratorie croniche (BPCO e asma) che condividono l'indicazione terapeutica per la maggior parte dei farmaci; l'unico farmaco specifico per la BPCO è il tiotropio, mentre per l'asma i farmaci specifici sono solo gli antagonisti dei recettori leucotrienici.

### **METODI**

- 1) Per stimare l'incremento di sensibilità della prevalenza di BPCO, sono stati definiti nuovi algoritmi che modificano il precedente per :
  - l'inclusione tra i casi prevalenti di BPCO, di pazienti dimessi con diagnosi di BPCO negli otto anni precedenti quello di stima, invece che nei 4 previsti dal precedente algoritmo;
  - l'inclusione tra i casi prevalenti, di casi BPCO definiti in base al consumo di farmaci. Due definizioni sono state adottate in alternativa; la prima è basata sull'uso di farmaci respiratori (codice ATC R03) ed è stata messa a punto dal CINECA<sup>4b</sup>; la seconda è basata sul consumo di farmaci per il trattamento della BPCO, come definito nel protocollo GOLD (disponibile al sito [www.goldcopd.com/](http://www.goldcopd.com/)).

I nuovi algoritmi vengono riportati in dettaglio nell'allegato 1.

I nuovi algoritmi sono stati testati per la stima di prevalenza della BPCO nel 2006, in tre città: Roma, Modena e Rovigo. La città di Roma era già inclusa nello studio precedente, mentre Modena e Rovigo sono coinvolte nello studio pilota del progetto CCM 2008 sui sistemi di sorveglianza epidemiologica basati sull'utilizzo di archivi elettronici sanitari.

Le modifiche agli algoritmi sono state apportate una alla volta e quindi, in combinazione. In tutti gli algoritmi si è mantenuta la selezione dei soggetti di età maggiore a 34 anni.

La costruzione degli algoritmi ha seguito le modalità riportate nel supplemento di E&P 2008: ai casi individuati dai dati SDO con approccio longitudinale a 5 o a 9 anni, sono stati aggiunti i casi identificati dal registro CM che la fonte ospedaliera non aveva visto; quindi sono stati aggiunti i casi identificati dalla farmaceutica, non visti dalle fonti precedenti.

- 2) La sensibilità e la specificità della definizione di caso BPCO da fonte ospedaliera, sono state stimate mediante il confronto con i casi BPCO definiti dal consumo farmaci.

I casi BPCO definiti dal consumo di farmaci sono stati distinti in certi e dubbi in accordo al rotocollo CINECA, e in casi lievi, moderati e gravi in accordo al protocollo GOLD.

Per la stima della sensibilità abbiamo assunto che il consumo continuativo di farmaci rappresentasse un criterio di conferma del caso ospedalizzato; mentre la specificità dei casi ospedalieri come casi medio-gravi di BPCO è stata stimata mediante il confronto con i casi gravi di BPCO definiti dal consumo di farmaci, rispetto ai casi lievi.

- 3) La capacità di discriminare i casi di BPCO individuati dal consumo dei farmaci rispetto a quelli di asma, è stata analizzata stimando il numero di pazienti dimessi con diagnosi di asma, che hanno consumato farmaci per la BPCO e che non hanno avuto diagnosi di BPCO.

## **RISULTATI**

La tabella 1 mostra una stima di prevalenza della BPCO più elevata sia per l'algoritmo che utilizza 9 anni di osservazione dei casi ospedalieri che per l'algoritmo che utilizza i dati della farmaceutica, rispetto all'algoritmo di base. Le stime più elevate si ottengono includendo entrambe le modifiche nello stesso algoritmo, ma gli incrementi più importanti sono dovuti comunque alla fonte farmaci. Per questa ragione, abbiamo deciso di utilizzare gli algoritmi 2 e 3 nelle analisi successive.

Le stime di prevalenza della BPCO mostrano valori più elevati a Roma rispetto a Modena e Rovigo (tabella 2). Il contributo dei casi con BPCO confermata dal consumo continuativo di farmaci, che non vengono ricoverati nei 5 precedenti è mediamente pari al numero di pazienti ospedalizzati per BPCO nello stesso periodo, in tutte le città: 56% a Roma, 49% a Modena e 50% a Rovigo.

Le tabelle 3 e 4 mostrano i risultati delle analisi di sensibilità e specificità della definizione di caso BPCO da fonte ospedaliera. Abbiamo già osservato che i casi ospedalieri di BPCO rappresentano il 50% circa dei casi prevalenti; se assumiamo che solo i casi che consumano in modo continuativo farmaci respiratori dopo la dimissione siano BPCO confermate, il contributo dei casi ospedalieri ai casi prevalenti di BPCO confermata si riduce al 18% a Roma, al 29% a Modena e al 28% a Rovigo (tabella 3), perché solo un terzo dei casi ricoverati per BPCO a Roma e a Modena risponde ai criteri di conferma; a Rovigo i casi confermati sono un po' di più (40%).

Al contrario, se limitiamo l'analisi ai pazienti dimessi con diagnosi di BPCO e confermati dal consumo continuativo di farmaci dopo la dimissione, i casi ospedalizzati sono casi certi di BPCO con il 94% di probabilità e sono molto pochi i casi ricoverati con una BPCO dubbia (tabella 4).

La tabella 5 mostra infine i risultati della possibile misclassificazione asma/BPCO. Un po' più della metà dei casi dimessi con diagnosi di asma consuma farmaci secondo lo schema terapeutico che identifica i casi di BPCO, tuttavia i casi di asma inclusi tra i casi di BPCO definiti dal consumo di farmaci sono circa il 3% in tutte e tre le città. Va sottolineato comunque che l'uso della diagnosi ospedaliera di asma misconosce i casi di asma non ricoverati che possono essere erroneamente classificati come casi di BPCO. Questi casi potrebbero essere in percentuali più elevate del 3% osservato.

## **CONCLUSIONI**

1. Gli algoritmi che utilizzano i farmaci aumentano la sensibilità della stima di prevalenza della BPCO da 1,5 a 2 volte rispetto all'algoritmo di base, sia con la definizione CINECA che con la definizione GOLD. La fonte farmaci fornisce un contributo pari a quello della fonte ospedaliera alla stima dei casi prevalenti di BPCO. Questo supporta l'ipotesi di una sottostima di circa il 50% dei casi prevalenti basata sul precedente algoritmo..
2. Solo un terzo dei pazienti BPCO ospedalizzati è confermato dal consumo continuativo di farmaci dopo la dimissione. Al contrario, tra i pazienti con BPCO confermata dall'uso di farmaci, le BPCO ospedaliere sono forme più gravi di malattia con una alta specificità. Questi dati contraddicono l'ipotesi della selezione dei casi gravi tra quelli ospedalizzati. I casi ricoverati probabilmente includono sia pazienti che necessitano di un inquadramento diagnostico per sintomi iniziali sia casi gravi tra quelli già confermati.
3. La inclusione dei casi di asma tra quelli di BPCO basata solo sul consumo dei farmaci è stata stimata inferiore al 3%. Tuttavia questa stima non tiene conto dei casi di asma non ospedalizzati.

## **RACCOMANDAZIONI**

Quando la prevalenza di BPCO è stimata in base ai dati delle ospedalizzazioni e dei decessi, si deve tenere conto che le stime possono essere dimezzate rispetto alla reale dimensione del fenomeno.

I casi di BPCO identificati per avere una diagnosi clinica e per il consumo continuativo di farmaci hanno una elevata probabilità di essere casi veri di BPCO di gravità medio-alta; questo criterio sembra essere quello più affidabile per la selezione di coorti BPCO.

**Tabella 1. Stime di prevalenza della BPCO per l'anno 2006.  
Confronti tra algoritmi di base e nuovi algoritmi**

<i>Algoritmi di base</i>	tassi di prevalenza % residenti di 35+ anni, standardizzati per età								
	Roma <sup>1</sup>			Modena			Rovigo		
	%	IC 95%		%	IC 95%		%	IC 95%	
<b>SDO trasv+CM</b>	1,08	1,06	1,09				0,66	0,61	0,70
<b>SDO long5+CM</b>	3,89	3,86	3,92	1,86	1,82	1,90	2,08	2,00	2,16
<i>Nuovi algoritmi:</i>									
<i>modifiche di un solo componente per volta rispetto a SDO long5+CM</i>									
algoritmo 1									
<b>SDO long9+CM</b>	5,59	5,56	5,62				2,94	2,84	3,04
algoritmo 2									
<b>SDO long5+CM+ farmaci CINECA</b>	6,35	6,50	6,57	3,77	3,71	3,82	4,31	4,19	4,43
algoritmo 3									
<b>SDO long5+CM+ farmaci GOLD</b>	5,77	5,73	5,80				3,83	3,72	3,94
<i>modifiche di più componenti in sequenza rispetto a SDO long5+CM</i>									
algoritmo 1 e 2 combinati									
<b>SDO long5+CM</b>	3,89	3,86	3,92				2,08	2,00	2,16
<b>SDO long9+CM</b>	5,59	5,6	5,6				2,94	2,84	3,04
<b>SDO long9+CM+ farmaci CINECA</b>	8,07	8,03	8,11				4,97	4,85	5,10

Algoritmo n.1 e 3 combinati

---

<b>SDO long5+CM</b>	3,89	3,86	3,92
<b>SDO long9+CM</b>	5,59	5,6	5,6
<b>SDO long9+CM+ farmaci GOLD</b>	7,37	7,3	7,4

---

Note

SDO = schede dimissione ospedaliera; CM = registro delle cause di morte

Il contributo esclusivo dei farmaci è stato stimato per l'anno 2006 senza effettuare il follow-up individuale di 365 giorni dalla data del primo acquisto

<sup>1</sup>la residenza per i dati di Roma è quella riportata dai singoli archivi sanitari elettronici, senza controllo anagrafico.

**Tabella 2. Stime di prevalenza della BPCO nell'anno 2006 ottenute dai dati di dimissione ospedaliera, di mortalità e di consumo farmaci (due definizioni).**

	tassi di prevalenza % residenti di 35+ anni, stardardizzati per età											
	Roma			Modena			Rovigo					
	n. casi	%	IC 95%	n. casi	%	IC 95%	n. casi	%	IC 95%			
<b>SDO long5+CM</b>	51734	3,17	3,14 3,20	8276	1,86	1,82 1,90	2608	2,08	2,00 2,16			
<b>SDO long5+CM+ farmaci CINECA</b>	117430	7,23	7,20 7,27	16369	3,77	3,71 3,82	5215	4,31	4,19 4,43			
<b>SDO long5+CM+ farmaci GOLD</b>	73983	4,56	4,53 4,59				4652	3,83	3,72 3,94			

**Note**

SDO = schede dimissione ospedaliera; CM = registro delle cause di morte

Le stime di prevalenza sono relative ai residenti, dopo verifica con i dati dell'anagrafe comunale o sanitaria per tutti i comuni.

Il contributo esclusivo dei farmaci è stato stimato per l'anno 2006 a Modena e Rovigo; mentre a Roma è stato introdotto il follow-up individuale di 365 giorni dalla data del primo acquisto, quando questo è avvenuto nel 2006. Il follow-up si è concluso nel 2007 per la metà dei soggetti circa.

Il dettaglio del contributo assoluto ed esclusivo delle varie fonti, per città, è riportato foglio successivo.

**Tabella 3. Sensibilità delle stime di casi prevalenti di BPCO da fonte SDO-CM rispetto ai casi prevalenti da fonte farmaceutica (farmaci respiratori R03; protocollo CINECA)**

casi BPCO fonte SDO long5 + CM	Roma			Modena			Rovigo			
	casi BPCO fonte farmaci			casi BPCO fonte farmaci			casi BPCO fonte farmaci			
	si	no	tot	si	no	tot	si	no		
si	14172	37562	51734	si	3240	5036	8276	si	1009	1599
no	65696	??	??	no	8093	??	??	no	2607	??
tot	79868	??		tot	11333	??		tot	3616	??
	sensibilità		17.7%	sensibilità		28.6%	sensibilità			
	vpp		27.4%	vpp		39.1%	vpp			

**Note**

Il numero di casi che consumano farmaci è stato stimato mediante un follow-up.

Il follow-up parte dal 1-1-2006 per i casi ricoverati negli anni precedenti e per quelli arruolati dal registro CM.

Il follow-up parte dalla data di dimissione per i casi ricoverati nell'anno di stima.

Per i casi da fonte farmaci (contributo esclusivo), il follow up parte dal 1-1-2006 a Rovigo e Modena; mentre a Roma è stato introdotto il follow-up individuale di 365 giorni dalla data del primo acquisto, quando questo è avvenuto nel 2006.

Le stime sono relative ai residenti, dopo verifica con i dati dell'anagrafe comunale o sanitaria per tutti i comuni.

**Tabella 4. Specificità delle stime dei casi prevalenti di BPCO da fonte SDO-CM rispetto ai casi certi di BPCO da fonte farmaceutica tra i soggetti che consumano farmaci respiratori (R03 Protocollo CINECA).**

	Roma			Modena			Rovigo				
	casi BPCO fonte farmaci			casi BPCO fonte farmaci			casi BPCO fonte farmaci				
	certi	dubbi	tot	certi	dubbi	tot	certi	dubbi			
casi BPCO fonte SDO long5 + CM	si	13375	797	14172	si	3030	210	3240	si	959	50
	no	56069	14792	65696	no	6658	1435	8093	no	2286	321
	tot	69444	15589	79868	tot	9670	1639	11333	tot	3233	370
	sensibilità		19.3%	sensibilità		31.2%	sensibilità				
	vpp		94.4%	vpp		93,7%	vpp				
	specificità		94.9%	specificità		87,6%	specificità				

**Note**

Il numero di casi che consumano farmaci è stato stimato mediante un follow-up.

Il follow-up parte dal 1-1- 2006 per i casi ricoverati negli anni precedenti e per quelli arruolati dal registro CM.

Il follow-up parte dalla data di dimissione per i casi ricoverati nell'anno di stima.

Per i casi da fonte farmaci (contributo esclusivo), il follow-up parte dal 1-1-2006 a Rovigo e Modena; mentre a Roma è stato introdotto il follow-up individuale di 365 giorni dalla data del primo acquisto, quando questo è avvenuto nel 2006.

Le stime sono relative ai residenti, dopo verifica con i dati dell'anagrafe comunale o sanitaria per tutti i comuni.

**Tabella 5. Capacità di discriminare i casi di BPCO dai casi di asma, per i soggetti di 35+ anni, quando si utilizza la fonte farmaci (protocollo CINECA)**

	Roma <sup>1</sup>			Modena			Rovigo					
	casi asma <sup>2</sup>			casi asma <sup>2</sup>			casi asma <sup>2</sup>					
	fonte SDO long5			fonte SDO long5			fonte SDO long5					
	si	no	tot	si	no	tot	si	no	tot			
<b>casi BPCO fonte farmaci (contributo esclusivo)</b>	si	2147 45,4%	83626 97,4%	85773 100%	si	341 45,3%	10992 97,0%	11333 100%	si	95 47,5%	3521 97,4%	3616 100%
	no	2584 54,6%	? <sup>3</sup>		no	412 54,7%	?		no	105 52,5%	? <sup>3</sup>	
	tot	4731 100%			tot	753			tot	200 100%		

**Note**

<sup>1</sup> la residenza per i dati di Roma è quella riportata negli archivi sanitari elettronici, senza controllo anagrafico.

<sup>2</sup> pazienti di 35+ anni d'età, dimessi con diagnosi di asma sia in principale che in secondaria, arruolati mediante approccio longitudinale a 5 anni, per i quali non è stata riportata la BPCO, né come diagnosi principale né tra le diagnosi secondarie nei 5 anni precedenti.

<sup>3</sup> questa casella dovrebbe riportare il numero di pazienti che non consuma farmaci respiratori secondo il protocollo CINECA e non ha avuto diagnosi ospedaliera di asma nei 5 anni precedenti. Questa frequenza non è stata calcolata, anche se avrebbe permesso una stima della specificità della definizione di caso BPCO rispetto ai casi di asma.

**Allegato D**  
**Algoritmo n.1**

<b>Fonte dei dati e criteri di definizione di caso prevalente di BPCO</b>		
<b>Fonte</b>	<b>Criteri selezione</b>	<b>Criteri definizione caso prevalente</b>
SDO	essere stati dimessi con diagnosi principale o secondaria di BPCO, definita mediante codici ICD-9 CM 490 - 492, 494, 496	essere stati dimessi con diagnosi di BPCO o essere deceduti per BPCO nell'anno di interesse  oppure
CM	essere deceduti per una causa identificata mediante codici ICD-9 CM 490 - 492, 494, 496	<b>essere stati dimessi con diagnosi di BPCO in uno degli 8 anni precedenti se vivi al 1 gennaio dell'anno di interesse</b>

**Algoritmo n.2**

<b>Fonte dei dati e criteri di definizione di caso prevalente di BPCO</b>		
<b>Fonte</b>	<b>Criteri selezione</b>	<b>Criteri definizione caso prevalente</b>
SDO	essere stati dimessi con diagnosi principale o secondaria di BPCO, definita mediante codici ICD-9 CM 490 - 492, 494, 496	essere stati dimessi con diagnosi di BPCO o essere deceduti per BPCO nell'anno di interesse  oppure
CM	essere deceduti per una causa identificata mediante codici ICD-9 CM 490 - 492, 494, 496	essere stati dimessi con diagnosi di BPCO in uno dei 4 anni precedenti se vivi al 1 gennaio dell'anno di interesse
<b>Farmaci</b>	<b>avere acquistato farmaci con codice ATC R03)</b>	oppure  <b>aver consumato farmaci respiratori, in quantità e tempi<sup>1</sup> tali da supportare la cronicità della patologia respiratoria.</b>

<sup>1</sup>almeno 5 confezioni di un farmaco respiratorio in 4 mesi (pazienti cronici certi), oppure 3-10 confezioni in 2-3 mesi o 3-4 confezioni in 4-6 mesi in monoterapia (pazienti cronici dubbi).

**Algoritmo n.3**

<b>Fonte dei dati e criteri di definizione di caso prevalente di BPCO</b>		
<b>Fonte</b>	<b>Criteri selezione</b>	<b>Criteri definizione caso prevalente</b>
SDO	essere stati dimessi con diagnosi principale o secondaria di BPCO, definita mediante codici ICD-9 CM 490 - 492, 494, 496	essere stati dimessi con diagnosi di BPCO o essere deceduti per BPCO nell'anno di interesse  oppure
CM	essere deceduti per una causa identificata mediante codici ICD-9 CM 490 - 492, 494, 496	essere stati dimessi con diagnosi di BPCO in uno dei 4 anni precedenti se vivi al 1 gennaio dell'anno di interesse
<b>Farmaci</b>	<b>Avere acquistato farmaci per la BPCO, tra quelli con codice ATC R03AC, R03BB, R03DA, R03BA, R03AK.</b>	oppure  <b>aver consumato farmaci per BPCO secondo il protocollo GOLD, sia in mono che poli-terapia, in quantità tali<sup>2</sup> da supportare la diagnosi di BPCO.</b>

<sup>2</sup> almeno 5 confezioni di un farmaco incluso nei 5 gruppi ATC sopra riportati in 1 anno (pazienti cronici certi), oppure 3-4 confezioni degli stessi farmaci in 1 anno (pazienti dubbi) sia in mono che in poli-terapia.

### 3. NUOVO ALGORITMO PER LA INDIVIDUAZIONE DEI DISTURBI ALLA GHIANDOLA TIROIDE: PAZIENTI CON IPER- E IPO-TIROIDISMO.<sup>5</sup>

#### Premessa

Da diversi anni la occorrenza di una patologia a carico della ghiandola tiroidea sembra essere in continuo aumento in tutta la popolazione ,ma non e' possibile quantificarne il reale impatto sulla popolazione generale priva di selezione. Tale aumento, che può essere dovuto in parte ad una maggiore sensibilità diagnostica dei medici e in parte ad un reale aumento, non e' dunque facilmente quantificabile.

Partendo quindi dai flussi integrati di ricoveri ospedalieri e prescrizioni farmaceutiche abbiamo provato a costruire un algoritmo per la individuazione dei casi di Iper e Ipo tiroidismo, tra i residenti della città di Taranto negli anni 2004-2007.

#### Materiali

Come definito nel progetto precedente e' stato effettuato un linkage individuale tra anagrafe assistiti e schede di dimissione ospedaliera (SDO, inclusa la mobilità passiva) con 4 chiavi: nome, cognome, data di nascita, comune di nascita, e che si e' svolto in due fasi:

Fase 1: linkage automatico con corrispondenza esatta delle 4 chiavi,

Fase 2: linkage decisionale con risoluzione manuale dei soggetti non linkati in fase 1.

Il programma di linkage e' stato messo a punto presso l'Istituto di Fisiologia Clinica del CNR di Pisa.

Sono state selezionate tutte le prescrizioni farmaceutiche (PF) relative ai farmaci usati in caso di iper-tiroidismo (ATC\_3: H03B\*) e ipotiroidismo (ATC\_5: H03AA\*).

Per l'arruolamento dei casi di ipotiroidismo e' stato anche usato, l'archivio delle esenzioni ticket per Ipotiroidismo acquisito ( codice 027.244) che era disponibile al momento dell'analisi.

#### Metodi

Come nel precedente progetto e' stato individuato il primo evento nel periodo dalle varie fonti, e quindi sono stati integrati i flussi per definire i contributi esclusivi e assoluti.

#### Definizione dei contributi esclusivi e assoluti

Per ogni categoria diagnostica è stato calcolato il contributo di ciascuna fonte alla definizione dei casi, distinguendo un «contributo esclusivo», determinato da soggetti identificati come malati per mezzo di quell'unica fonte, e «contributo assoluto», come numero totale di casi alla cui identificazione ha contribuito la fonte considerata. I contributi esclusivi e assoluti sono stati calcolati stratificando per classi di età, sesso e periodo.

#### Definizione di prevalenza di periodo e incidenza annuale

Per ogni individuo sono stati individuati nel periodo esaminato il primo ricovero, la prima prescrizione farmaceutica.

E' stata quindi calcolata la **prevalenza di periodo** come numero di casi rilevati tramite il contributo esclusivo dalle varie fonti usate nel periodo dal 2004 al 2007 e la **incidenza annuale** nei due anni 2006 e 2007 come numero di casi, che non abbiano avuto un ricovero o una prescrizione negli anni precedenti a partire dal 2004.

---

<sup>5</sup> A cura di Maria Angela Vigotti (Universita',Pisa e IFC-CNR,Pisa), Giorgio Iervasi (IFC-CNR,Pisa), Anna Romanelli (IFC-CNR,Pisa), Mauro Raciti(IFC-CNR,Pisa), Giuseppe Leone (ASL-TA), Sante Minerba (ASL-TA)

La prevalenza di periodo e' stato calcolata rispetto alla popolazione media residente, calcolata come media di quella riportata dall'ISTAT al 1 gennaio di ogni anno e la incidenza annuale rispetto alla popolazione residente al 1 gennaio di quell'anno.<sup>6</sup>

Sia per la prevalenza di periodo che per l'incidenza annuale sono stati calcolati i tassi grezzi e standardizzati per età x100 residenti, secondo il metodo diretto, usando come riferimento la popolazione italiana totale al censimento 2001.

### 3.1. IPER-TIROIDISMO

#### Selezione della Casistica

##### Schede di Dimissione Ospedaliera.

Sono stati selezionati tutti i ricoveri con diagnosi di dimissione principale o secondaria per Tireotossicosi (ICD.9=242). Sono stati individuati 717 ricoveri: 215 tra gli uomini e 502 tra le donne.

##### Prescrizioni farmaceutiche.

Sono state selezionate tutte le prescrizioni farmaceutiche con codice ATC\_3: H03B\* (Preparati antitiroidei), che includono in codice ATC\_5, che riguardano esclusivamente il farmaco Metazole.

PF per iper-tiroidismo. Taranto 2004-2007		
Codici ATC_5	Descrizione	
H03BA	Tiouracilici	-
H03BB	Derivati imidazolici contenenti zolfo (metazole)	8760
H03BC	Perclorati	7
H03BX	Altri preparati antitiroidei	2
	Totale	8769

#### Esenzioni tickets

Fonte dei dati e criteri di definizione di caso incidente di IPER_ TIROIDISMO		
Fonte	Criteri di selezione	Criteri per la definizione di incidenza
SDO	Ricovero per Tireotossicosi con o senza gozzo (ICD-9-CM: 242) in diagnosi di dimissione principale o secondaria	sono state considerate le fonti SDO e PF; Assenza di prescrizioni e di altri ricoveri, con diagnosi di dimissione principale o secondaria
PF	Almeno una prescrizione Preparati antitiroidei (Codice ATC_3: H03B*)	con i codici ICD-9-CM 242 ( Tireotossicosi con o senza gozzo ), almeno per 2 anni precedenti

<sup>6</sup> [http://demo.istat.it/pop2004\(2005-2006-2007\)/index.html](http://demo.istat.it/pop2004(2005-2006-2007)/index.html)

## Risultati

### Primo ricovero.

Dal file dei 717 ricoveri nel periodo 2004-2007 e' stato individuato il primo ricovero per ogni individuo: su 581 primi ricoveri (81.8%) riguardano le donne.

Inoltre i primi ricoveri, in base alle diagnosi principale o secondarie, sono stati assegnati, in modo sequenziale, alle categorie di patologie indicate nella tabella seguente.

Primo ricovero per causa e sesso. Taranto 2004-2007						
Codice Gruppo	Descrizione Gruppo	ICD.9	Descrizione	Uomini	Donne	Totale
1	Gozzo tossico	242.0	Gozzo tossico diffuso	28	73	101
2	Gozzo nodulare	242.1 242.2 242.3	Gozzo tossico uninodulare Gozzo multinodulare tossico Gozzo nodulare tossico non specificato	71	196	267
3	Tireotossicosi	242.4 242.8 242.9	Tireotossicosi da tessuto tiroideo ectopico Tireotossicosi di altra origine specificata Tireotossicosi senza menzione di gozzo o di altra causa	70	143	213
			totale	169	412	581

### Prima prescrizione farmaceutica

Negli anni 2004-2007, tramite la prima prescrizione farmaceutica nel periodo, sono stati individuati 1602 individui che assumono preparati antitiroidei.

### Contributi esclusivi e assoluti

Nella tabella sono descritti i casi individuati in base al contributo di ogni fonte.

Sono stati individuati 18490 nuovi casi di ipertiroidismo nel periodo 2004-2007, di cui l'73.5% sono donne(N=1360).

Ipertiroidismo							
Contributi delle varie fonti nella individuazione di nuovi casi nel periodo 2004-2007							
		Totale N	Uomini N	Donne N	Totale %	Uomini %	Donne %
	totale nuovi casi	1849	489	1360	100.0	100.0	100.0
Contributi Esclusivi	Solo Ricoveri	247	58	189	13.4	11.9	13.9
	Solo Farmaci	1268	320	948	68.6	65.4	69.7
	Ric.+Farm.	334	111	223	18.1	22.7	16.4
Contributi assoluti	Ricoveri	581	169	412	31.4	34.6	30.3
	Farmaci	1602	431	1171	86.6	88.1	86.1

Il maggiore contributo alla identificazione dei casi e' dato dai farmaci che individuano nel loro complesso circa l'87% dei pazienti.

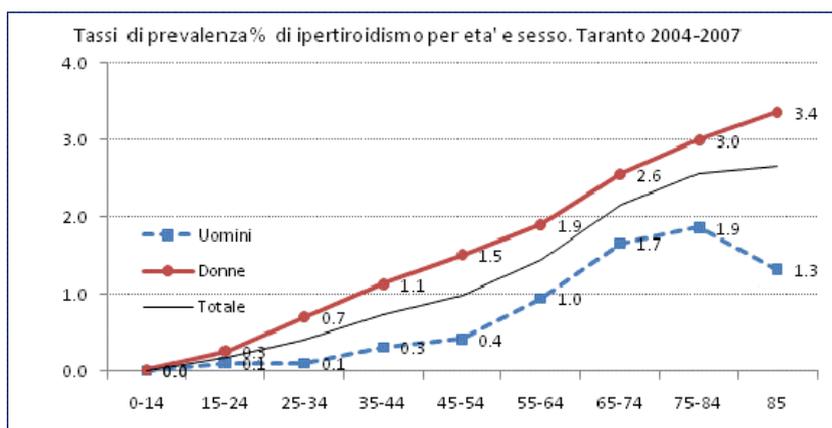
## Prevalenza

Il tasso standardizzato di prevalenza % di ipertiroidismo nel periodo 2004-2007 (numero di casi sulla popolazione media nel periodo), qui di seguito riportati, indicano valori di oltre 2 volte superiori tra le donne che tra gli uomini.

Tassi di prevalenza % di ipotiroidismo. Taranto 2004-2007				
	N.casi	TG	TSD	LC95%
Uomini	489	0.518	0.564	0.559 ; 0.569
Donne	1360	1.313	1.265	1.258 ; 1.272
Totale	1849	0.934	0.944	0.940 ; 0.948

TG= tasso grezzo; TSD= tasso standardizzato

Nella figura successiva sono riportati i tassi di prevalenza specifici per classe di età, da cui si evidenzia un andamento diversificato nei due sessi. Tra le donne la patologia è presente sin dalle età più giovani con valori che aumentano in modo costante fino alle età più anziane. Negli uomini i valori sono inferiori a quelli delle donne e cominciano ad aumentare intorno ai 60.



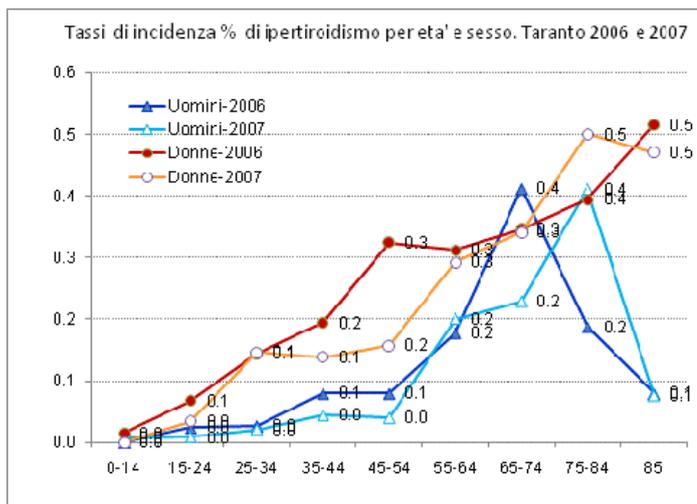
## Incidenza

Per gli anni 2006 e 2007 sono stati calcolati i tassi standardizzati di incidenza %, per i nuovi casi, che quindi non erano presenti negli anni precedenti (a partire dal 2004), e qui di seguito riportati.

Tassi di incidenza % di ipertiroidismo per anno. Taranto 2006-2007									
	2006				2007				
	N.	TG	TSD	LC95%	N.	TG	TSD	LC95%	
Uomini	95	0.101	0.107	0.105 ; 0.109	81	0.087	0.092	0.090 ; 0.094	
Donne	228	0.220	0.214	0.211 ; 0.216	196	0.190	0.180	0.177 ; 0.182	
Totale	323	0.163	0.164	0.162 ; 0.166	277	0.141	0.139	0.138 ; 0.141	

TG= tasso grezzo; TSD= tasso standardizzato

Nella figura sono riportati i tassi di incidenza specifici per classe di età, da cui si conferma un lieve decremento nel 2007; si conferma l'andamento crescente sin dalle età più giovani tra le donne e sopra i 60 anni negli uomini.



### 3.2. IPO – TIROIDISMO

#### Selezione della Casistica

#### Schede di Dimissione Ospedaliera.

Sono stati selezionati tutti i ricoveri con le cause indicate nella tabella presenti nelle diagnosi di dimissione principale o secondarie.

Numero di ricoveri per causa e sesso. Taranto 2004-2007					
Ordine di selezione	Codice ICD.9	Descrizione cause	Uomini	Donne	Totale
1	244.0-244.1	Ipotiroidismo post-chirurgico (§)	49	189	238
2	243, 246.1	Ipotiroidismo congenito e Gozzo disormonogenico	5	28	33
3	245.2-245.9	Tiroidite	136	815	951
4	244.2-244.9	Ipotiroidismo acquisito	258	1025	1283
5	193,226	Tumore alla tiroide maligno o benigno (*)	2	18	20
		totale	450	2075	2525

(§) inclusi i casi con menzione di ipertiroidismo precedente  
 (\*) Individuati tra i casi selezionati per le cause su indicate, vedi testo

I ricoveri sono stati selezionati nell'ordine indicato nella tabella. Tra i ricoveri selezionati sono stati quindi identificati i ricoveri con indicazione, in diagnosi principale o secondaria, di tumore maligno (ICD.9=193) o benigno (ICD.9=226). In totale sono stati selezionati 2525 ricoveri nel periodo (2075 donne e 450 uomini).

#### Prescrizioni farmaceutiche.

Sono state selezionate tutte le prescrizioni farmaceutiche con codice ATC\_5: H03AA\* (ormoni tiroidei). Le PF riguardano essenzialmente la levotiroxina sodica, come risulta dalla tabella seguente.

<b>PF per ipo-tiroidismo. Taranto 2004-2007</b>		
<b>Codici ATC_7</b>	<b>Descrizione</b>	
H03AA01	levotiroxina sodica	136825
H03AA02	liotironina sodica	615
H03AA03	associazioni di levotiroxina e liotironina	0
H03AA04	tiratricolo	0
H03AA05	preparati a base di ghiandola tiroide	0
	totale	137440

### **Esenzioni ticket**

I codici esenzione che riguardano l'ipotiroidismo sono riportati nella tabella.

Al momento dell'analisi per i residenti a Taranto erano disponibili solo le esenzioni per ipotiroidismo acquisito (027-244), per un totale di 884 esenzioni di cui 120 uomini e 764 donne.

Inoltre nei dati disponibili non era riportato l'anno in cui è stata data la esenzione, per cui non e' stato possibile assegnare l'anno preciso né calcolare l'età.

<b>Esenzioni per ipo-tiroidismo.</b>	
<b>Codice</b>	<b>Descrizione</b>
056-245.2	tiroidite di Hashimoto
027-244	ipotiroidismo acquisito
027-243	ipotiroidismo congenito
048-193	tumori maligni della ghiandola tiroide

<b>Fonte dei dati e criteri di definizione di caso incidente di IPO_ TIROIDISMO</b>		
<b>Fonte</b>	<b>Criteri selezione casistica</b>	<b>Criteri definizione incidenza</b>
le fonti e i criteri SDO	Ricovero per una delle cause scelte in diagnosi di dimissione principale o secondaria	Stato considerato le fonti SDO, PF, ET, riportati nello schema seguente: Assenza di prescrizioni e di altri ricoveri ,con diagnosi di dimissione principale o secondaria
PF	Prescrizione per Ormoni tiroidei: Codice ATC_5 H03AA*	con i codici ICD-9-CM 242 ( Tireotossicosi con o senza gozzo ), almeno per 2 anni precedenti
ET	Esenzione tickets per Ipotiroidismo acquisito (grave): Codice 027-244	

## **Risultati**

### **Primo ricovero.**

Dal file dei 2525 ricoveri nel periodo 2004-2007 e' stato individuato il primo ricovero per ogni individuo: su 1946 primi ricoveri 1592 (81.8%) riguardano le donne.

Inoltre i primi ricoveri, in base alle diagnosi principale o secondarie, sono stati assegnati, in modo sequenziale, alle categorie di patologie indicate nella tabella.

Primo ricovero per causa e sesso. Taranto 2004-2007				
ICD.9	Descrizione	Uomini	Donne	Totale
244.0-244.1	Ipotiroidismo post-chirurgico(§)	39	160	200
243, 246.1	Ipotiroidismo congenito e Gozzo disormonogenico	5	16	21
245.2-245.9	Tiroidite	117	627	744
244.2-244.9	Ipotiroidismo acquisito	191	776	967
193,226	Tumore alla tiroide maligno o benigno (*)	2	12	14
	totale	354	1592	1946

(§) inclusi i casi con menzione di ipertiroidismo precedente  
 (\*) Individuati tra i casi selezionati per le cause su indicate, vedi testo

### Prima prescrizione farmaceutica

Negli anni 2004-2007, tramite la prima prescrizione farmaceutica nel periodo, sono stati individuati 10624 individui che assumono farmaci ipotiroidici.

### Contributi esclusivi e assoluti

Nella tabella sono descritti i casi individuati in base al contributo di ogni fonte.

Sono stati individuati 10810 nuovi casi di ipotiroidismo nel periodo 2004-2007 di cui l'81.1% sono donne(N=8766).

Ipotiroidismo							
Contributi delle varie fonti nella individuazione di nuovi casi nel periodo 2004-2007							
		Totale	Uomini	Donne	Totale	Uomini	Donne
		N	N	N	%	%	%
	totale nuovi casi	10810	2044	8766	100.0	100.0	100.0
Contributi Esclusivi	Solo Ricoveri	491	113	378	4.5	5.5	4.3
	Solo Farmaci	8165	1596	6569	75.5	78.1	74.9
	Solo Esenzione	51	10	41	0.5	0.5	0.5
	Ric.+Farm.	1270	215	1055	11.7	10.5	12.0
	Es.+Farm.	648	84	564	6.0	4.1	6.4
	Es.+Ric.	4	2	2	0.0	0.1	0.0
	Es.+Ric.+Farm.	181	24	157	1.7	1.2	1.8
Contributi assoluti	Ricoveri	1946	354	1592	18.0	17.3	18.2
	Farmaci	10264	1919	8345	94.9	93.9	95.2
	Esenzioni	884	120	764	8.2	5.9	8.7

Il maggiore contributo alla identificazione dei casi è dato dai farmaci che individuano nel loro complesso circa il 95% dei pazienti.

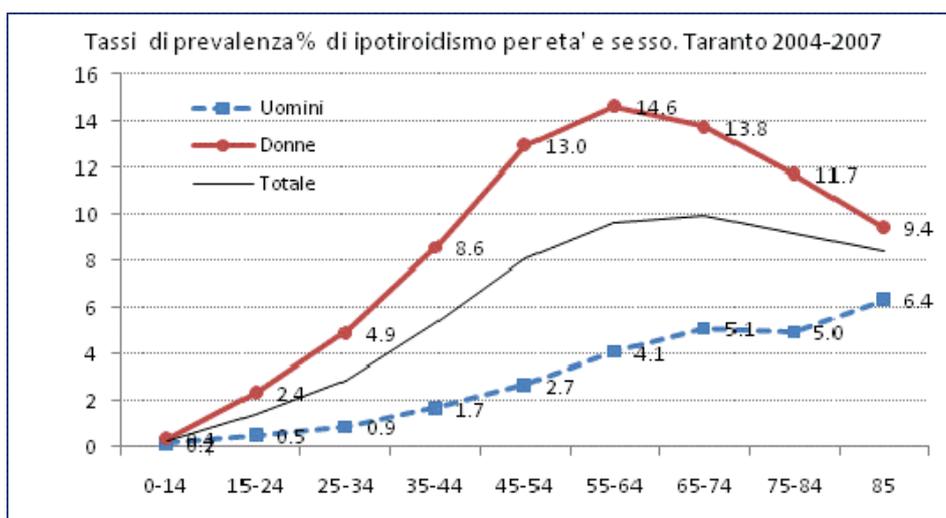
### Prevalenza

I tassi standardizzati di prevalenza % di ipotiroidismo nel periodo 2004-2007 (numero di casi sulla popolazione media nel periodo), qui di seguito riportati, indicano valori di oltre 3 volte superiori tra le donne rispetto a gli uomini.

Tassi di prevalenza % di ipotiroidismo. Taranto 2004-2007 (*)				
	N.casi	TG	TSD	LC95%
Uomini	2034	2.15	2.30	2.29 ; 2.31
Donne	8725	8.42	8.20	8.19 ; 8.22
Totale	10759	5.43	5.45	

(\*) esclusi i 51 casi identificati solo con le esenzioni ; TG= tasso grezzo; TSD= tasso standardizzato

Nella figura successiva sono riportati i tassi di prevalenza specifici per classe di età, da cui si evidenzia un andamento diversificato nei due sessi. Tra le donne la patologia è presente sin dalle età più giovani con valori che aumentano in modo rapido fino ai 35-44 anni e subiscono un calo dopo i 55-64 anni. Negli uomini i valori sono molto inferiori a quelli femminili e non subiscono un calo con l'età, l'andamento nelle età più giovani indica una crescita più rapida fino ai 35-44 anni.



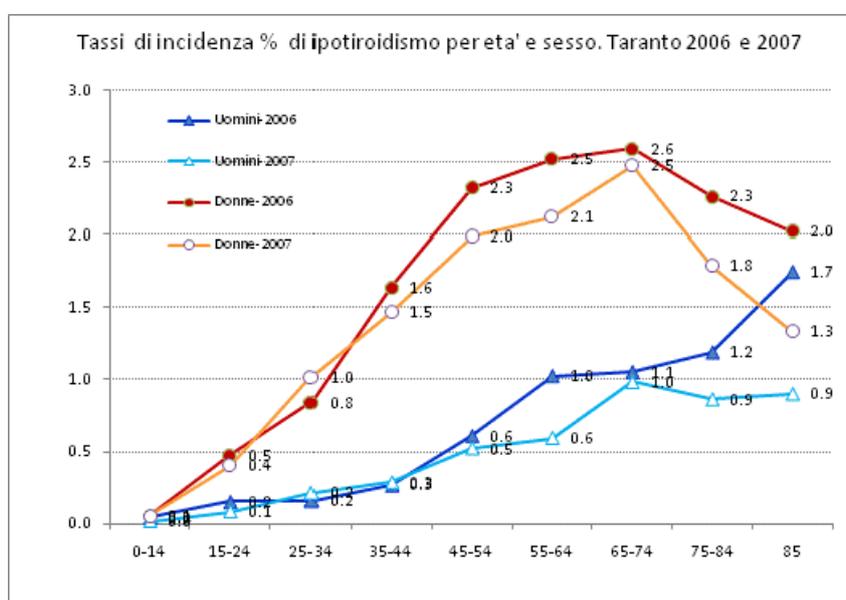
## Incidenza

Per gli anni 2006 e 2007 sono stati calcolati i tassi standardizzati di incidenza %, per i nuovi casi, che quindi non erano presenti negli anni precedenti (a partire dal 2004), e qui di seguito riportati.

Tassi di incidenza % di ipotiroidismo per anno. Taranto 2006-2007(*)									
	2006				2007				
	N.	TG	TSD	LC95%	N.	TG	TSD	LC95%	
Uomini	455	0.483	0.513	0.508 ; 0.518	359	0.384	0.402	0.398 ; 0.407	
Donne	1601	1.548	1.503	1.495 ; 1.510	1432	1.391	1.349	1.342 ; 1.356	
Totale	2056	1.041	1.041	1.036 ; 1.045	1791	0.912	0.906	0.902 ; 0.910	

(\*) esclusi i 51 casi identificati solo con le esenzioni ; TG= tasso grezzo; TSD= tasso standardizzato

Nella figura sono riportati i tassi di incidenza specifici per classe di età, da cui si osserva nel 2007 una riduzione dei nuovi casi nelle età superiori ai 35-44 anni tra le donne e superiori ai 45-54 anni tra gli uomini. Nelle età più giovani non si osserva riduzione.



## Conclusioni

I dati di prevalenza riportati sono confrontabili con quelli riportati nella letteratura internazionale<sup>7, 8</sup>. La numerosità rilevata dai flussi sanitari verosimilmente sottostima la incidenza di tali patologie nella popolazione in quanto rileva chi effettivamente si cura (stima basata cioè sul consumo del farmaco) e/o chi ha una diagnosi che per sua importanza viene riportata nella diagnosi finale di ricovero. L'ipotiroidismo invece è noto che, anche se diagnosticato, frequentemente non viene trattato specie nell'anziano fino a che non diventa conclamato.

La bontà di tali analisi è confermata dalle differenze osservate fra i due sessi che concordano con quanto riportato in letteratura e soprattutto dalle differenze per fascia di età con anticipo della diagnosi e cura nelle donne rispetto agli uomini, come atteso da un punto di vista clinico.

Da tali considerazioni quindi si può ritenere che l'algoritmo utilizzato sia valido per lo scopo che ci eravamo proposti.

<sup>7</sup> Biondi B Cooper DS The clinical significance of subclinical thyroid dysfunction. *Endocr Rev* 2008 29(19 76-131)

<sup>8</sup> Roberts CG, Ladenson PW Hypothyroidism. *Lancet* 2004 6 (363) 793-803

## 4. MONITORAGGIO DEL PERCORSO ASSISTENZIALE DEL DIABETE <sup>9</sup>

### **Premessa**

La rilevanza del diabete quale problema di sanità pubblica è ben nota. L'elevato impatto della malattia sulla salute della popolazione e sulle risorse sanitarie fa sì che gli studi epidemiologici volti a monitorare la frequenza di malattia, la sua evoluzione e la qualità dell'assistenza erogata rivestano un grande interesse sia per i clinici che per gli amministratori della sanità. Il Piano Nazionale della Prevenzione 2004-2006 e le sue successive proroghe hanno recepito questa esigenza, individuando come obiettivi prioritari «l'attivazione di un sistema di monitoraggio della malattia su base informatizzata » e «identificare i pazienti diabetici attraverso l'iscrizione in appositi registri».

La disponibilità di sistemi di sorveglianza per monitorare la malattia in diverse aree geografiche e in diversi sottogruppi di popolazione è di grande rilievo. Archivi di soggetti con diabete raccolti per fini amministrativi (ricoveri ospedalieri, prescrizioni farmaceutiche, esenzioni ticket, prestazioni ambulatoriali, mortalità) sono ormai facilmente disponibili in tutte le ASL e possono essere utilizzati sia per la ricostruzione di popolazioni di diabetici, sia per monitorarne il processo assistenziale e gli esiti.

L'obiettivo di questo lavoro è sperimentare le possibilità di monitoraggio del percorso assistenziale di persone con diabete, attraverso l'utilizzo di sistemi informativi correnti.

### **Metodi**

La realizzazione del progetto si è articolata in alcuni passaggi successivi, che hanno visto il coinvolgimento di unità operative differenti:

1. individuazione della popolazione di diabetici residenti;
2. follow-up di indicatori di processo del percorso assistenziale;
3. follow-up di indicatori di esito del percorso assistenziale.

### **1. La popolazione diabetica**

Per l'identificazione della popolazione con diabete è stato utilizzato il protocollo messo a punto dal gruppo di lavoro dell'Associazione Italiana di Epidemiologia (AIE) con la Società Italiana di Statistica Medica ed Epidemiologia Clinica (SISMEC), già applicato in tre grandi aree metropolitane (Torino, Firenze, Venezia) e una ASL del Veneto (Alto Vicentino).

Per una dettagliata descrizione dei metodi utilizzati a questo scopo si rimanda al lavoro pubblicato<sup>10</sup>. La tabella 1 sintetizza l'algoritmo utilizzato per l'identificazione dei casi.

---

<sup>9</sup> A cura di Roberto Gnani, Roberta Picariello, Alessandro Barchielli, Cristina Canova

<sup>10</sup> *Epidemiol Prev* 2008; 32(3) suppl 1:15-21

**Tabella 1. Fonti dei dati e criteri di definizione di caso prevalente di diabete**

Fonte	Criteri selezione casistica	Criteri definizione prevalenza
SDO	Dimissione nell'anno di stima o nei 4 anni precedenti con diagnosi di diabete (ICD9-CM = 250) in qualunque campo diagnosi e non deceduto al 1 gennaio dell'anno di stima	Presenza in almeno una delle fonti
Prescrizioni farmaceutiche	Almeno due prescrizione nell'anno di stima per farmaci antidiabetici (ATC = A10*)	
Esenzione ticket	Codice di esenzione 013.250 rilasciato nell'anno di stima o nei tre anni precedenti e non deceduto al 1 gennaio dell'anno di stima	

## **2. Follow up di indicatori di processo del processo assistenziale**

Questa parte dello studio ha interessato i residenti della città di Torino, dell'ASL 10 di Firenze e dell'AULSS 18 di Rovigo.

Per ognuno dei centri è stata individuata la popolazione con diabete mediante l'applicazione dell'algoritmo di cui al punto 1, alla data del 31 luglio 2003 (Torino), del 31 dicembre 2003 (Firenze) e del 1 gennaio 2005 (Rovigo).

Le popolazioni così definite sono state appaiate mediante codice identificativo univoco (codice fiscale) agli archivi delle prestazioni specialistiche ambulatoriali, allo scopo di individuare le prestazioni riportate nella tabella 2, effettuate nei 12 mesi successivi alla data indice. Nel caso di più di una prestazione dello stesso tipo è stata considerata la prestazione con la data più prossimo all'inizio del follow-up.

E' stato così possibile calcolare alcuni dei più comuni indicatori di aderenza alle linee guida raccomandate per il monitoraggio e l'assistenza ai malati di diabete.<sup>11,12,13,14</sup> La probabilità esatta di effettuazione di un esame o di una visita è stata stimata con metodo Kaplan-Meier.

<sup>11</sup> De Micheli A. Italian standards for diabetes mellitus 2007: executive summary Diabete Italia, AMD Associazione Medici

Diabetologi, SID Società Italiana di Diabetologia. Acta Diabetol 2008;45:107-127

<sup>12</sup> Italian Association of Clinical Diabetologists. *Quality Indicators in Diabetes in Italy*. AMD, Roma, 2008

<sup>13</sup> De Beaufort CE, et al. European Union diabetes indicators. *Eur J Public Health* 2003;13:51-54

<sup>14</sup> EuCID. *Final Report European Core Indicators in Diabetes Project, 2008*.

[http://ec.europa.eu/health/ph\\_projects/2005/action1/docs/action1\\_2005\\_frep\\_11\\_en.pdf](http://ec.europa.eu/health/ph_projects/2005/action1/docs/action1_2005_frep_11_en.pdf).

**Tabella 2. Prestazioni ambulatoriali e relativi codici utilizzati per**

<b>Prestazione</b>	<b>Codice prestazione</b>
Emoglobina glicata	prestazione <b>90.28.1</b>
Glicemia	Prestazione <b>90.27.1</b>
Colesterolo	prestazione <b>90.14.1</b> or <b>90.14.2</b> or prestazione <b>90.14.3</b>
Trigliceridi	Prestazione <b>90.43.2</b>
Microalbuminuria	prestazione <b>90.33.4</b>
Cretaininemia	prestazione <b>90.16.3</b>
Visita diabetologica	(prestazione <b>89.7</b> or <b>89.01</b> ) and (disciplina <b>17</b> or <b>19</b> )
Valutazione oculistica ( <i>visita oculistica o fundus oculi o retinografia</i> )	(prestazione <b>89.7</b> or <b>89.01</b> ) and (disciplina <b>34</b> ) or (prestazione <b>95.11</b> ) or (prestazione <b>95.09.1</b> )
Valutazione cardiologica ( <i>visita cardiologia o ECG</i> )	((prestazione <b>89.7</b> or <b>89.01</b> ) and (disciplina <b>8</b> )) or (prestazione <b>89.52</b> )

### **3. Follow up di indicatori di esito del percorso assistenziale**

Questa parte dello studio ha interessato i residenti della città di Torino e dell'AULSS 12 Veneziana.

Utilizzando le procedure descritte al punto 1 sono stati selezionati i diabetici vivi e residenti al 1 gennaio 2002 nei due centri. La restante popolazione residente è stata considerata come non diabetica. Sia la popolazione con diabete che la popolazione senza diabete è stata sottomessa a follow up di mortalità (totale e per causa) e di incidenza per Infarto Miocardico Acuto (IMA), ictus e amputazioni non traumatiche degli arti inferiori per un periodo di 4 anni (sino al 31 dicembre 2005).

Mediante *record-linkage* con gli archivi locali delle cause di morte sono stati selezionati i decessi (utilizzando la sola causa di morte principale) per malattie coronariche (ICD9 = 410–414) e vasculopatie cerebrali (ICD9 = 430-438).

I casi incidenti di IMA e di ictus sono stati identificati utilizzando l'algoritmo del gruppo di lavoro dell'Associazione Italiana di Epidemiologia (AIE) e Società Italiana di Statistica Medica ed Epidemiologia Clinica (SISMEC)<sup>15, 16</sup>, utilizzando gli archivi locali dei ricoveri ospedalieri e delle cause di morte.

Le amputazioni degli arti inferiori sono state identificate mediante *record linkage* con gli archivi dei ricoveri ospedalieri, selezionando le dimissioni con codici DRG 113, 114, o 285. In caso di più di un ricovero per i DRG selezionati è stato considerato il ricovero più prossimo alla data di inizio del follow-up. L'inizio del follow-up è stato quindi definito al 1 gennaio 2002 e terminato alla data di incidenza, di decesso, di emigrazione (solo nel caso

<sup>15</sup> Barchielli A, Balzi D, Bruni A, et al. Stima dell'incidenza dell'infarto miocardico acuto basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane. *Epidemiol Prev* 2008; 32 (Suppl 1):30-7

<sup>16</sup> Tancioni V, Collini F, Balzi D, Barchielli et al. Stima dell'incidenza di ictus miocardico acuto basata su dati sanitari correnti mediante un algoritmo comune in differenti aree italiane. *Epidemiol Prev* 2008; 32 (Suppl 1):38-45.

di Torino) o termine del follow-up (31 dicembre 2005). I soggetti emigrati da Torino nel periodo di follow-up sono stati considerati “persi al follow-up” (4.7% dei diabetici, and 8,6% dei non diabetici), mentre questa informazione non era disponibile per i residenti a Venezia, con conseguente sottostima dei tassi a Venezia. Per valutare l’effetto del diabete sugli eventi in studio sono stati calcolati i Rapporti Standardizzati di Mortalità (SMR) e di Incidenza (SIR), separatamente per uomini e donne, utilizzando come standard i tassi di mortalità/incidenza per fasce di età quinquennali della popolazione non diabetica.

## **Risultati**

### Il processo assistenziale

La tabella 3 riporta, nelle tre aree partecipanti, la probabilità di effettuare nell’arco di 12 mesi alcune prestazioni specialistiche raccomandate dalle linee guida per una buona assistenza dei pazienti diabetici. I livelli più elevati, per tutte le prestazioni, si evidenziano a Torino, mentre percentuali nettamente inferiori (mediamente dell’ordine del 50%) si registrano a Firenze; Rovigo si colloca costantemente in una posizione intermedia.

**Tabella 3. probabilità annua (in percentuale) di alcune prestazioni ambulatoriali in soggetti con diabete in tre città.**

<b>Prestazione</b>	<b>Firenze</b> (n=35.312)	<b>Torino</b> (n=33.453)	<b>Rovigo</b> (n=8.524)
Emoglobina glicata	38.1	71.0	56.7
Glicemia	56.1	79.8	68.7
Trigliceridi	34.6	62.8	42.3
Colesterolemia	35.5	64.7	44.0
Creatininemia	40.0	69.1	44.3
Microalbuminuria	14.8	31.0	5.9
Visita cardiologia	10.2	33.3	14.5
Visita diabetologica	5.0	67.5	48.8
Visita oculistica	6.9	23.6	4.3

### Gli esiti

Sono state identificate 44.889 persone con diabete (11.853 a Venezia e 33.036 a Torino) e 950.229 senza diabete (216.402 a Venezia e 733.827 a Torino) di età superiore a 20 anni al 1 gennaio 2002. I diabetici sono stati seguiti per 160.679 anni-persona per mortalità, per 158.795 anni-persona per incidenza di IMA, 159.672 anni-persona per incidenza di ictus and 160.228 per le amputazioni. Nei 4 anni di follow-up si sono registrati 7.599 decessi tra i diabetici e 41.966 tra i non diabetici, 1.620 e 7.302 IMA, 1.461 e 7.821 ictus, 319 e 236 amputazioni.

La tabella 4 riporta, nelle due aree di interesse, gli SMR e i SIR. I risultati sono pressoché

sovrapponibili nelle due città. Il rischio di morte tra i diabetici è più elevato di circa il 70-75% rispetto ai non diabetici in entrambi i sessi; i rischi più elevati sono per cardiopatia ischemica. Tra i diabetici, rispetto ai non diabetici, l'incidenza di IMA è più elevata di oltre due volte per gli uomini e di oltre due volte e mezzo per le donne, l'incidenza di ictus è doppia in entrambi i sessi, mentre il rischio di amputazione è da 10 a 20 volte superiore.

**Tabella 3. SMR e SIR dei diabetici confrontati con non diabetici per sesso a Torino e Venezia; 2002-2005.**

<u>Città</u> <b>Causa morte /incidenza</b>	<b>Uomini</b>			<b>Donne</b>		
	n	SMR/SIR	(95% IC)	n	SMR/SIR	(95% IC)
<b>Venezia</b>						
<b><u>Mortalità</u></b>						
<i>Totale</i>	1182	174.7	(164.9-185.0)	1046	175.6	(165.1-186.5)
<i>Cardiopatia schemica</i>	234	214.7	(188.1-244.1)	201	219.2	(189.9-251.7)
<i>Vasculopatia cerebrale</i>	94	162.0	(130.9-198.2)	116	156.4	(129.2-187.5)
<b><u>Incidenza</u></b>						
<i>IMA</i>	231	217.0	(189.9-246.8)	191	257.4	(222.2-296.6)
<i>Ictus</i>	201	190.57	(165.1-218.8)	200	188.69	(163.4-216.7)
<b><u>Amputazioni</u></b>	90	1733.6	(1394-2131)	50	2130.2	(1580-2808)
<b>Torino</b>						
<b><u>Mortalità</u></b>						
<i>Totale</i>	2755	167.7	(161.5-174.1)	2616	174.8	(168.2-181.7)
<i>Cardiopatia ischemica</i>	425	213.6	(193.8-234.9)	367	238.1	(214.3-263.7)
<i>Vasculopatia cerebrale</i>	263	156.3	(137.9-176.3)	321	129.0	(115.2-143.9)
<b><u>Incidenza</u></b>						
<i>IMA</i>	658	195.1	(180.5-210.6)	540	290.2	(266.3-315.8)
<i>Ictus</i>	525	182.0	(166.8-198.2)	535	176.7	(162.1-192.4)
<b><u>Amputazioni</u></b>	113	1022.1	(842-1229)	66	967.5	(748-1231)

### ***Discussione***

Questo lavoro ha sperimentato la fattibilità di un sistema di sorveglianza dell'assistenza ricevuta da pazienti con diabete mediante l'utilizzo di fonti informative correnti, comunemente disponibili presso le aziende sanitarie della maggior parte delle regioni italiane.

Il lavoro ha preso l'avio dalla definizione di una popolazione di residenti con diabete individuata secondo l'algoritmo già sperimentato in una precedente fase del progetto.

Per quanto riguarda il monitoraggio di indicatori di processo calcolati a partire dal *record*

*linkage* con l'archivio delle prestazioni ambulatoriali, sono stati evidenziati risultati molto differenti nelle tre aree di sperimentazione. Le differenze osservate possono essere attribuibili a reali differenze nella qualità dell'assistenza o a differenze nella completezza dell'archivio utilizzato. Nel primo caso verrebbe evidenziato un forte svantaggio nel follow up clinico dei diabetici residenti a Firenze, in particolare rispetto ai diabetici residenti a Torino (con Rovigo in una posizione intermedia); anche nel caso di Torino, comunque, va sottolineato come gli standard di qualità non siano ottimali, dal momento che le linee guida attuali indicano l'effettuazione di almeno 1 esame dell'emoglobina glicata all'anno (altre LG indicano un esame ogni 6 mesi). Nel secondo caso le differenze osservate potrebbero essere dovute a una differente completezza degli archivi delle prestazioni specialistiche o ad una differente qualità del *linkage*. Sono in corso valutazioni sugli archivi per accertarne la completezza e la qualità delle variabili utilizzate come chiave di *linkage*. Non vi sono invece, al momento, altri elementi di letteratura o empirici per confermare una reale differenza nella qualità dell'assistenza. La letteratura nazionale riporta risultati migliori di quelli registrati a Torino, ma riferiti a soli soggetti seguiti dai centri di diabetologia<sup>17</sup>, mentre a livello europeo la variabilità tra Paesi è molto ampia<sup>18</sup>, simile a quella rilevata nel nostro lavoro tra i centri partecipanti. A una differenza di monitoraggio clinico come quella riportata nel nostro lavoro potrebbe corrispondere una simile differenza in termini di esiti a distanza della malattia (sopravvivenza, incidenza di complicanze); al momento, non è stato possibile effettuare un follow up di mortalità/ricovero dei pazienti con diabete a Firenze e a Rovigo per valutare differenze negli esiti.

Follow up di mortalità e di incidenza sono invece state effettuati a Torino e a Venezia mediante *record linkage* con gli archivi di mortalità e di dimissioni ospedaliere. I risultati ottenuti nelle due città sono del tutto sovrapponibili tra loro e in linea con la letteratura corrente.

In conclusione:

1. Il monitoraggio del percorso assistenziale dei pazienti diabetici attraverso *record linkage* con l'archivio delle prescrizioni ambulatoriali è fattibile e di grande rilevanza in termini di sanità pubblica. E' possibile calcolare, con facilità e modesto impegno di risorse, numerosi indicatori di qualità del processo di assistenza del diabete tra quelli indicati dalle linee guida attuali. Il principale limite risiede nella qualità degli archivi delle prestazioni specialistiche. Trattandosi di archivi di grandi dimensioni da poco entrati nell'uso *routinario* delle Aziende è possibile che le procedure di acquisizione dei dati e di controlli di qualità non siano ancora consolidate in tutte le ASL. Si rende quindi indispensabile, prima di avviare un uso estensivo di questa procedura, un'attenta valutazione di qualità di questi archivi.

2. Il monitoraggio degli esiti a distanza dei pazienti diabetici mediante *record linkage* con archivi di mortalità e di ricoveri ospedalieri (singolarmente o integrati per l'individuazione di casi incidenti) è fattibile e di grande utilità in termini di sanità pubblica. La consistenza nei risultati ottenuti in due centri differenti e la coerenza con la letteratura esistente depone per una buona affidabilità dei risultati ottenuti.

3. L'integrazione delle due esperienze compiute (valutazione del processo e valutazione degli esiti) delinea la possibilità di realizzare, a costi limitati, un sistema di sorveglianza di popolazione (cioè non selezionato per afferenza a centri specialistici o su campioni di popolazione) dell'intero Percorso Diagnostico Terapeutico Assistenziale (PDTA) del diabete: dall'insorgenza della malattia, alla qualità dell'assistenza ricevuta, sino all'insorgenza di complicazioni o al decesso.

---

<sup>17</sup> Italian Association of Clinical Diabetologists. *Quality Indicators in Diabetes in Italy*. AMD, Roma, 2008.

<sup>18</sup> EuCID. *Final Report European Core Indicators in Diabetes Project*, 2008.

[http://ec.europa.eu/health/ph\\_projects/2005/action1/docs/action1\\_2005\\_frep\\_11\\_en.pdf](http://ec.europa.eu/health/ph_projects/2005/action1/docs/action1_2005_frep_11_en.pdf)

### **Area 3. Gruppo di lavoro VALIDAZIONE**

**Componenti: Pietro Barbieri, Cristina Canova, Piergiorgio Duca, Annunziata Faustini, Claudia Galassi, Mauro Maistrello, Enrica Migliore, Lorenzo Simonato, Mariangela Vigotti**

#### **PREMESSA**

L'utilizzo di database amministrativi per l'indagine epidemiologica, variamente finalizzata, è ormai molto comune, data la larga diffusione dei database, la loro completezza estensione e buona qualità di compilazione e registrazione.

Meno comune è la validazione dei risultati che tali database possono fornire al ricercatore epidemiologo, sia usati individualmente sia in combinazione fra loro, soprattutto se si tratta di ricavarne stime di prevalenza e/o incidenza di malattia nella popolazione generale.

Diverso è il caso dell'uso di database per identificare soggetti affetti da una certa patologia, per cui si richiede una elevatissima specificità senza troppo dare peso alla sensibilità, o quando si tratta di individuare casi di malattia definiti in modo operativo più che concettuale, come capita per studi intesi alla valutazione di procedure o tecnologie sanitarie. Anche in questo caso non è di interesse immediato la qualità della stima che si ottiene, in relazione ad una popolazione definita su base geografica, dei "casi di malattia" ma piuttosto di tutti i "casi di richiesta di assistenza".

#### **RICERCA DI LETTERATURA**

Per esplorare in che misura, e secondo quali modalità, la validazione di nostro interesse è stata affrontata da parte diversi ricercatori, il gruppo di lavoro ha proceduto anzitutto ad una ricerca bibliografica in PubMed, volta ad identificare i lavori intesi a validare algoritmi di identificazione di casistica che portassero a stime "accurate" di occorrenza, utilizzando i database amministrativi.

È stata per questo effettuata una ricerca in MEDLINE con la seguente strategia: ("epidemiology"[Subheading] OR "epidemiology"[All Fields] OR "epidemiology"[MeSH Terms]) AND administrative[all fields] AND databases [all fields]) che ha prodotto 674 risultati.

#### **SELEZIONE DELLA LETTERATURA**

Una prima selezione, effettuata in base ai titoli, ha portato a identificare 144 lavori di interesse in quanto:

1) mirati alla validazione di stime di occorrenza di patologie di interesse per il progetto di ricerca

2) di presumibile interesse metodologico pur finalizzati alla validazione su patologie non incluse fra quelle di interesse per la nostra ricerca

che sono stati valutati, a questo punto, in base al contenuto dell'abstract (risultato mancante in soli 4 casi). Ai 144 selezionati se ne sono 9 ricavati da segnalazioni estemporanee (referenze di letteratura) come documentato nella Bibliografia riportata alla fine.

La lettura degli abstract è stata effettuata individualmente o in doppio, distribuendoli in 7 blocchi, con il mandato a ciascun "valutatore" di classificare ogni lavoro in una di 3 categorie:

A: escluso

Indicare il motivo dell'esclusione

B: da segnalare ad altri gruppi  
archivio,  
validazione

Indicare il motivo della segnalazione (validazione  
modalità di gestione di archivio, algoritmo senza

per patologia di interesse o per altra comunque da

indicare)

C<sub>1</sub>: da selezionare in quanto orientato alla validazione di procedura di stima di prevalenza e/o incidenza di una o più delle patologie di nostro specifico interesse (Diabete – BPCO – IMA/SCA – patologia tiroide)

C<sub>2</sub>: da selezionare in quanto utile per gli aspetti metodologici che propone, valuta, applica per patologia di interesse o altra patologia.

62 sono stati i lavori esclusi in questa fase mentre i restanti sono stati sottoposti a lettura estesa.

## MODALITA' DI ESTRAZIONE DELL'INFORMAZIONE

La lettura è stata eseguita seguendo una griglia che prevedeva, per ogni lavoro, una breve sintesi (50 – 100 parole) con l'informazione essenziale circa:

- 1) Rationale
- 2) Popolazione
- 3) Materiali, strumenti e metodi
- 4) Risultati
- 5) Conclusioni degli autori

In particolare si richiedeva di considerare:

*Informazioni da estrarre:*

- 1) Rationale, obiettivi e contesto in cui si svolge lo studio (potrebbe anche essere un esempio di validazione all'interno di uno studio con altri obiettivi prioritari)
- 2) Popolazione (nazione, paese), dimensione numerica, periodo di esecuzione, setting, estensione geografica (locale, provinciale, regionale, nazionale, internazionale)
- 3) Tipo di studio (se applicabile: coorte storica, coorte prospettica, caso-controllo, casistica, trasversale)
- 4) Istituzioni coinvolte (ospedale, medicina di base, uffici governativi, uffici del servizio sanitario locale e/o centrale)
- 5) Data base utilizzato per estrazione casi, procedure/algoritmi utilizzati, strumenti
- 6) Reference standard utilizzato per la validazione, procedure/algoritmi utilizzati, strumenti
- 7) Modalità di validazione seguita, metodi statistici e software applicato
- 8) Risultati:  
Tipi di stima ottenuti: sensibilità, specificità, valori predittivi, riproducibilità  
Tipo di misure validate: incidenza/prevalenza  
Confronti fra metodi eventualmente effettuati, valutazioni di sensibilità  
Indicazione di costo e di fattibilità (generalizzabilità, riproducibilità)  
Confronti dei valori ottenuti per le misure a livello nazionale, internazionale, temporale
- 9) Conclusioni degli autori

e in conclusione riportare:

*Giudizi da formulare;*

- 1) Commenti del relatore sulla qualità del report:  
Archivi usati e procedure seguite descritti in modo esauriente  
Criteri di selezione/estrazione inclusione/esclusione descritti in modo esauriente  
Data elements e procedure di stima descritte in modo esauriente



29 Hansel A, Hollowell J, McNiece R, Nichols T, Strachan D	147
30 Merry AHH, Boer JMA, Leo J. Schouten	148
31 Dendukuri N, Joseph L	149
32 Qu Y, Tan M, Kutner MH	150
33 Koek HL, de Bruin A, Gast A, Gevers E, Kardaun JW, Reitsma JB, et al	151
34 Bernatsky S, Joseph L, Brelisle P, Boivin J-F, Rajan R, Moores A, Clarke A	153

(6)

#### RAZIONALE

Editoriale che introduce un intero numero di “Annali of Epidemiology” dedicato agli studi epidemiologici occupazionali basati sull’uso di svariate tipologie di database amministrativi, utilizzati con diverse finalità e variamente valutati per completezza, validità dei risultati forniti, semplicità di utilizzo.

#### OBIETTIVI

Non trattandosi di uno studio originale non fornisce risultati immediatamente utili o applicabili, ma si segnala per l’interesse che la consultazione dell’intero numero della rivista può rivestire per coloro che intendessero avere un primo approccio alla problematica.

In particolare si va dall’uso di database per accertare lo stato in vita degli appartenenti ad una specifica coorte occupazionale, a quello di un database appropriato alla identificazione di una serie di controlli utili per a condurre uno studio retrospettivo su base di popolazione, fino all’uso dei dati di mortalità per l’identificazione di pattern geografici di mortalità per tumori.

#### COMMENTI RELATORE

Si segnala per suggerire il fascicolo della rivista di cui funge da introduzione e traccia di lettura.

(12)

#### OBIETTIVI

Validare l'occorrenza di alcune patologie croniche: diabete, ipertensione, BPCO e ulcera gastroduodenale come stimata mediante intervista per l'Indagine Stato di Salute (ISS) ISTAT usando i dati di un database della SIMG (550 MMG)

#### POPOLAZIONE IN STUDIO

Italia. Due campioni di popolazione di età > 14 anni; il primo di 199799 individui è stato intervistato tra il luglio 1999 e il giugno 2000 dall' ISTAT; il secondo di 432747 è stato campionato dal DB-SIMG tra i pazienti vivi alla fine del 2000

PERIODO IN STUDIO : anno 2000

DATABASE : risultati della VI indagine sullo stato di salute ISTAT e database MMG (solo SIMG)

#### DEFINIZIONE DI CASO

- nell'indagine di popolazione ISTAT è basata sulle risposte a due domande principali: "E' stato mai affetto da una di queste malattie?" e "Le è mai stata diagnosticata una di queste patologie da un medico ?"
- nel database dei MMG : presenza della diagnosi (codici ICD-9) e almeno una prescrizione del gruppo generale di farmaci. Ad es. per la BPCO uno dei farmaci respiratori identificati dal codice ATC R03.

MISURE DI OCCORRENZA : prevalenza

STANDARD PER LA VALIDAZIONE : prevalenza stimata mediante i dati del DB MMG

MODALITÀ DI VALIDAZIONE : confronto dei tassi di prevalenza standardizzati per età e correlazione tra età e differenze nella prevalenza tra ISTAT e DB MMG

#### RISULTATI

Le prevalenze standardizzate per età del diabete e dell'ipertensione sono confrontabili per gli uomini, mentre per le donne le stime sono più alte di circa il 10% quando le patologie sono auto riferite.

Le prevalenze standardizzate per età della BPCO e dell'ulcera gastroduodenale sono più elevate sia negli uomini che nelle donne quando le patologie sono autoriferite.

La correlazione tra aumento dell'età e differenze della prevalenza è rilevante nelle donne con frequenze di malattia più elevate nell'indagine mediante intervista, con valori elevati per il diabete ( $R^2 = 0.82$ ,  $p=0.001$ ) e ipertensione ( $R^2=0.44$ ,  $p=0.05$ ).

#### CONCLUSIONI DEGLI AUTORI

Esistono differenze nelle stime dei due DB per patologia, per sesso ed età dei pazienti. Per diabete e ipertensione, patologie di facile diagnosi e con elevato ricorso alle prestazioni del medico curante, il DB-MMG è meno costoso. Per BPCO e ulcera è possibile una minore validità (difetto o eccesso di registrazione) nel DB-MMG.

#### COMMENTI RELATORE

La qualità del lavoro lascia molto a desiderare: poco è detto sulla rappresentatività del campione dei MMG, discutibile è il metodo di confronto dei risultati, sorprendente è la conclusione in cui si inverte l'obiettivo e si valuta il DB-MMG rispetto all'indagine di popolazione ISTAT.

(18)

#### RAZIONALE

L'identificazione di coorti di diabetici è strumento necessario alla valutazione di esito e di performance (processo) delle attività di prevenzione.

#### OBIETTIVI

Assegnare la validità di diverse strategie per identificare i casi prevalenti di diabete in popolazione.

#### TIPO DI STUDIO

Gli autori hanno condotto una revisione (non sistematica) della letteratura selezionando i lavori che stimano i casi prevalenti di diabete e che forniscono stime dirette o dati sufficienti alla stima di validazione dei casi. (Periodo di pubblicazione 1966 – 2002)

DATABASE differenti sono state valutate per la loro capacità di identificare i casi prevalenti di diabete: a) dati anagrafici relativi alla causa di morte; b) dati amministrativi quali i rimborsi per l'assistenza medica, i ricoveri, i dati delle prestazioni mediche, i dati della farmaceutica; c) dati di indagini campionarie di popolazione con interviste sulla diagnosi di diabete da parte di un medico.

DEFINIZIONE DI CASO Non sono riportate in dettaglio le definizioni adottate negli studi considerati.

#### MISURE DI OCCORRENZA Prevalenza

#### STANDARD PER LA VALIDAZIONE

Come gold standard sono state utilizzate le informazioni derivate dalle cartelle cliniche ospedaliere o del medico curante, oppure le interviste ai pazienti.

#### MODALITÀ DI VALIDAZIONE

Sensibilità, specificità, VPP e kappa sono state scelte come misure di validazione, mentre le misure di riproducibilità sono state escluse come insufficienti a validare i casi.

#### RISULTATI

Per i 17 lavori che hanno utilizzato i dati delle cause di morte è stata stimata solo la sensibilità, che mostra valori bassi (mediana di 36%; range 7.1% - 68.6%). Stime più alte (mediana 43%) sono riportate negli studi che utilizzano sia la causa principale che la causa concorrente di morte e in quelli che utilizzano dati di popolazione rispetto ai registri di patologia.

Per i 16 lavori che hanno utilizzato i dati amministrativi è stato possibile stimare tutti i parametri. La sensibilità, la specificità e il VPP presentano valori di 81.5%, 99% e 92%, rispettivamente e il Kappa è in un range di 67% - 96%. L'uso della diagnosi principale di dimissione ha una sensibilità più bassa rispetto all'uso della diagnosi in qualunque posizione, ma riduce anche la specificità.

Infine gli 11 lavori che hanno utilizzato l'informazione rilevata mediante interviste in indagini campionarie, ha presentato stime elevate di sensibilità, VPP, specificità e kappa con valori mediani di 81%, 89%, 99% e 80%, rispettivamente.

#### CONCLUSIONI DEGLI AUTORI

I registri di mortalità sono meno sensibili; i dati amministrativi danno risultati migliori, con buoni parametri di validazione. E' possibile su questa base sia stimare la prevalenza del diabete sia selezionare coorti per valutazioni di esito e di qualità dell'assistenza.

#### COMMENTI RELATORE

Metodi di validazione formalmente corretti. Manca il confronto tra le stime di prevalenza nei diversi studi. E' impossibile il confronto tra i diversi gold standard usati per la validazione, in quanto non sono applicati sulla stessa popolazione.

(33)

#### RAZIONALE

Nonostante l'utilizzo dei DB amministrativi sia prassi diffusa e presenti vantaggi per le stime di occorrenza in epidemiologia, è necessario validare la diagnosi delle singole patologie.

#### OBIETTIVI

Validare la prevalenza della BPCO stimata da database amministrativi.

#### POPOLAZIONE IN STUDIO

860705 soggetti di 65+ anni, residenti in Quebec, provincia del Canada, con assicurazione sanitaria.

PERIODO IN STUDIO 1994 -1999.

#### DATABASE

Registro generale delle assicurazioni mediche del Quebec: un database amministrativo dei rimborsi prestazioni mediche e farmaceutiche.

DEFINIZIONE DI CASO: un caso di BPCO "possibile" è stato definito come un paziente di 65+ anni, con almeno tre prestazioni per problemi relativi a BPCO (e mai per asma), che abbia avuto prescrizioni terapeutiche per ipratropio bromuro o  $\beta$ 2-agonisti. La presenza di una diagnosi di BPCO da parte di una specialista definiva il caso come confermato.

#### MISURE DI OCCORRENZA

Prevalenza

#### STANDARD PER LA VALIDAZIONE

Prevalenza stimata nella indagine di popolazione sullo stato di salute 1994/1995.

Analisi di sensibilità per i diversi codici ICD-9 che identificano la BPCO.

#### MODALITÀ DI VALIDAZIONE

Confronto tra le stime di prevalenza ottenute dai DB con quelle riportate nella Indagine di Popolazione. La validazione interna è stata valutata con analisi di sensibilità per età, definizioni basate su codici ICD-9 per la BPCO differenti (490, 491, 492, 496 versus 491, 492, 496) essendo il codice 490 il meno specifico.

#### RISULTATI

116225 pazienti di 65+ anni sono stati identificati nei DB come affetti da COPD, con una stima di prevalenza di BPCO del 13.5%, che mostra un incremento proporzionale all'età (10.7 % nei pazienti di 65-74 anni e 17.8% in quelli di età > ai 74 anni. I 2/3 dei pazienti BPCO hanno codici ICD-9 = 490; la loro proporzione è inversamente proporzionale all'età dei pazienti. le "BPCO possibili" sono il 37%, le probabili sono il 23% rispetto a tutti i pazienti per cui è riportata la BPCO almeno una volta.

La stima di prevalenza in popolazione è pari a 5.4% e 8.3% per i pazienti di 64-75 anni e di età > 74 anni, rispettivamente.

#### CONCLUSIONI DEGLI AUTORI

Gli autori concludono che la prevalenza di BPCO stimata dai dati del database dei rimborsi delle prestazioni non è valida.

#### COMMENTI RELATORE

Il lavoro non utilizza gold standard di validazione; ma rappresenta un buon esempio di come la validazione interna possa essere utilizzata correttamente per validare le stime. Apprezzabile in particolare la definizione di caso "a priori" che utilizzano gli autori.

(41)

Ampia monografia che riferisce i risultati di una revisione sistematica di letteratura che affronta metodicamente vantaggi, costi, limiti e utilità dell'uso dei database amministrativi soprattutto per finalità di valutazione di tecnologie sanitarie.

Si segnala soprattutto a quanti siano impegnati nel settore dell'epidemiologia valutativa, senza trascurare la rilevanza che nella monografia rivestono alcuni aspetti metodologici di valutazione critica dei database e del loro uso, che vengono affrontati dagli autori in modo sistematico.

(42)

#### RAZIONALE

È possibile usare dati di ricovero ospedaliero e di dimissione, insieme ai rimborsi medici e alle prescrizioni di farmaci per valutare la performance del sistema sanitario canadese.

#### OBIETTIVI

Determinare il grado di utilizzo dei servizi della British Columbia da parte di malati residenti nella provincia di Alberta, al confine.

#### METODI

Si sono stimati tassi di procedure cardiologiche utilizzando due tipi di registro clinico unitamente ai dati di ammissione/dimissione ospedaliera.

#### RISULTATI

Il fenomeno dei trasferimenti dei pazienti al di fuori dei confini regionali comporta un sistematico errore di stima nei tassi delle procedure sanitarie.

#### STIME

Tassi standardizzati per età e sesso di procedure cardiologiche (Caterizzazioni cardiace, Interventui coronarici percutanei, Bypass aorto-coronarici).

(61)

#### RAZIONALE

Valutare sensibilità e specificità nella identificazione di pazienti ipertesi

#### OBIETTIVI

Valutazione di accuratezza di due algoritmi (uno basato sulla diagnosi medica e l'altra su diagnosi medica più prescrizione farmaceutica) per la identificazione di pazienti ipertesi utilizzando come standard di riferimento i record clinici.

#### DATABASE

Sono state usate le registrazioni (richieste di rimborso per prestazioni) relative a trattamenti in regime di ricovero (inpatient), ambulatoriale (outpatient) e di prescrizione farmaceutica dal gennaio 2000 al marzo 2003.

#### RISULTATI

Fra 60.033 positivi e 68.725 negativi per l'algoritmo A e 1298 positivi e 73.460 negativi per l'algoritmo B furono identificate coppie di pazienti appaiati per caratteristiche cliniche e per "propensity score" da cui, potendo disporre per tutti dei dati estratti da record medici, sono stati selezionate 258 coppie per valutare l'algoritmo A e 138 per quello B.

#### STIME

Sono state stimate Sensibilità 70.8% e 76.2% e Specificità 74.9% e 93.3% per A e B rispettivamente. È stata calcolata anche una statistica Kappa di concordanza aggiustata per il caso che è risultata pari a 0.45 per A e 0.65 per B.

(67)

#### OBIETTIVI

Stimare incidenza e prevalenza di Lupus Eritematoso Sistemico usando sorgenti multiple di dati amministrativi di popolazione.

#### METODI

Utilizzo dei database di ospedalizzazione e di richiesta rimborso di prestazione medica per la provincia del Quebec nel periodo 1994 – 2003 con una copertura di circa 7.5 milioni di abitanti.

Sono stati applicati metodi di regressione (classi latenti) per valutare l'effetto su sensibilità e specificità di identificazione da parte di caratteristiche dei pazienti e dei medici.

#### RISULTATI

Si ottennero dati di prevalenza ( $38 \times 10^{-5}$ ) e di incidenza ( $3 \times 10^{-5}$ ) confrontabili sia utilizzando i dati di ospedalizzazione che i dati di rimborso, ma solo una ridotta proporzione di casi veniva identificata secondo entrambe le procedure.

#### CONCLUSIONI

Nel caso del LES la mancanza di un Gold Standard rende necessario utilizzare modelli statistici per ottenere stime di sensibilità e specificità.

(89)

#### RAZIONALE

I dati amministrativi possono essere usati per valutare il ricorso da parte dei pazienti ai servizi sanitari.

#### OBIETTIVI

Valutare l'accuratezza nella identificazione di casi di Morbo di Parkinson usando dati amministrativi e prendendo come gold standard le diagnosi autoriferite da parte dei pazienti.

#### METODI

I casi furono identificati a partire dal registro delle prescrizioni farmaceutiche dal 1992 al 2000 (72.922 prescrizioni per 30.469 pazienti).

#### RISULTATI

Vennero confrontate diverse stime di sensibilità e di valore predittivo positivo ottenute con regole di selezione differenti.

(100)

#### OBIETTIVO

Costruire e validare algoritmi per l'accertamento di casi di osteoporosi a partire da database amministrativi al fine di poter stimare la prevalenza della condizione nella popolazione utilizzando come gold standard i dati ricavati da un programma regionale di screening basato sulla densitometria minerale ossea.

#### METODI

Per la popolazione residente nel Manitoba (Canada) sono stati elaborati dati estratti da database ospedalieri, di prestazioni specialistiche e di prescrizioni farmaceutiche utilizzando metodi di classificazione basati su reti neurali, procedura di classificazione ad albero, reversione logistica.

#### RISULTATI

Le stime basate su database multipli risultarono più accurate.

(116)

#### RAZIONALE

È possibile stimare l'incidenza di cancro mammario a partire da database amministrativi ?

#### METODO

Si è proceduto a stimare la relazione fra incidenza ottenuta dai registri presenti in alcuni dipartimenti e i dati di ospedalizzazione così da poter applicare la relazione stimata nei dipartimenti nei quali non esiste una registrazione sistematica dei casi diagnosticati.

#### RISULTATI

I dati di incidenza furono ottenuti per 12 dipartimenti con registro per l'anno 2002 e venne studiata la relazione con il numero di ricoveri identificati solo in base alla diagnosi principale (algoritmo 1) o in base anche alla menzione dell'avvenuta resezione (algoritmo 2).

Mentre l'algoritmo 1 ha fornito stime più vicine alla verità, per i dipartimenti con registro, l'algoritmo 2 ha fornito stime più eterogenee fra dipartimenti.

(128)

#### RAZIONALE

La validazione delle definizioni di caso dai database (DB) richiede la verifica della diagnosi mediante surveys o revisioni della documentazione clinica. Questi interventi oltre ad essere impegnativi e costosi, non sono esenti da errori di misura che causano distorsioni sia della stima di prevalenza e della stima di sensibilità e specificità dei DB.

#### OBIETTIVI

1) sviluppare un modello bayesiano per validare i casi di asma come individuati dai DB amministrativi, 2) usare il modello per creare e validare una definizione di asma trattata, 3) stimare la prevalenza dell'asma nella popolazione in studio, 4) stimare sensibilità e specificità dei DB rispetto al gold standard del modello bayesiano..

#### POPOLAZIONE IN STUDIO

Canada - British Columbia

PERIODO IN STUDIO 1991-2001

#### DATABASE

- National Committee for Quality Assurance 2001 (definizione di caso: asma persistente)
- Il registro provinciale dell'asma (definizione di caso : 2 visite o un ricovero per asma o 3 prescrizioni per asma)
- Il DB dello studio Breathing easier (definizione di caso: una diagnosi clinica o 3 prescrizioni)

#### DEFINIZIONE DI CASO

Definizione di base un soggetto tra i 5 e i 55 anni (tra il 1996 e il 2001) registrato dal 1991 al 2001 in un DB con almeno un codice di prestazione ambulatoriale o di diagnosi di dimissione per faringite, bronchite, asma o BPCO.

Le 3 definizioni di caso complete per ciascun DB sono :

- asma persistente (MH);
- asma come accettata nel registro dell'asma (AR);
- la definizione dello studio Breathing easier (SS).

#### MISURE DI OCCORRENZA E DI VALIDAZIONE

Prevalenza di asma,

#### STANDARD PER LA VALIDAZIONE

Casi selezionati dal modello bayesiano a due classi (asma trattata e non asma) che utilizza anche i dati della farmaceutica.

#### MODALITÀ DI VALIDAZIONE

Confronto tra le definizioni di caso dei DB, deterministiche e non modificabili nel tempo, con quella del modello bayesiano Latent class model (LCM) a due classi, usato come gold standard.

#### RISULTATI

E' stato scelto come gold standard il modelli a due classi (asma trattata e non asma) in quanto risultava più stabile rispetto a quelli con tre o quattro classi, ovvero la

assegnazione dei pazienti alle due classi era più riproducibile per campioni diversi e in anni diversi.

Il confronto tra le definizioni di caso deterministiche dei 3 DB esaminati con il gold standard ha permesso di individuare

- il modello SS come quello più sensibile,
- i modelli MH e AR come quelli più specifici
- il modello SS come quello con la concordanza più alta rispetto al gold standard.

#### CONCLUSIONI DEGLI AUTORI

L'uso del modello proposto (LCM) per sviluppare una definizione di caso per l'asma trattata è utile

ad identificare una coorte di pazienti asmatici di popolazione.

(132)

#### OBIETTIVI

Il lavoro è finalizzato a stimare la prevalenza di asma e BPCO combinate, mediante il consumo di farmaci inalatori.

#### POPOLAZIONE IN STUDIO

Svezia, dimensione nazionale, soggetti di età superiore ai 20 anni. Nr. abitanti = 6756453.

PERIODO IN STUDIO anno 2004.

#### DATABASE

due database della farmaceutica; il primo riporta dati aggregati per età e sesso dei pazienti, per tipo e dosi giornaliere del farmaco. Il riporta le stesse informazioni a livello nominativo. I database sono stati confrontati in una contea.

#### DEFINIZIONE DI CASO

Non è riportata esplicitamente; ma dalla descrizione delle fonti sembra di poter sostenere che sia il numero di pazienti cui è stato prescritto almeno uno dei farmaci con indicazione per asma o BPCO.

MISURE DI OCCORRENZA prevalenza di malattie polmonari cronico-ostruttive (MPCO) che includono sia l'asma che la BPCO.

#### STANDARD PER LA VALIDAZIONE

Stime di precedenti survey, che riportano una prevalenza di 4-6% per la BPCO e dell'8% per l'asma.

#### MODALITÀ DI VALIDAZIONE

confronto con i dati di survey precedenti condotte in Svezia. Stime di correlazione tra i due registri per alcuni farmaci respiratori: broncodilatatori  $\beta$ 2 agonisti (ATC = R03AC) e muscarinici (ATC = R03BB), cortisonici inalatori (ATC = R03BA); formulazioni combinate di  $\beta$ 2 agonisti e cortisonici (ATC = R03AK).

#### RISULTATI

La prevalenza di MPCO, nel 2004 in Svezia è stata stimata pari a 6.37%.

I coefficienti di correlazione tra i registri farmaci sono di 0.974 per i broncodilatatori  $\beta$ 2 agonisti; 0.938 per le formulazioni combinate di  $\beta$ 2 agonisti e cortisonici; 0.991 per i cortisonici inalatori; 0.994 per i broncodilatatori antimuscarinici.

#### CONCLUSIONI DEGLI AUTORI

I risultati della prevalenza della MPCO stimati mediante il consumo dei farmaci sono confrontabili con le stime di prevalenza delle survey precedenti, che utilizzano altre fonti, quali i registri della medicina di base.

## COMMENTI RELATORE

Il lavoro non risponde ai due problemi più importanti nell'uso della fonte farmaci per le malattie respiratorie, cioè 1) se il consumo di farmaci sia una fonte esaustiva rispetto ai dati ospedalieri o dei MMB; 2) se la fonte farmaci consente di distinguere l'asma dalla BPCO.

(142)

## OBIETTIVI

confrontare l'accuratezza di due algoritmi per identificare il morbo di Parkinson, utilizzando i dati di DB amministrativi. Un algoritmo definisce i casi come pazienti che avessero il codice 332.0 almeno una volta in entrambi i setting (ambulatoriale ed ospedaliero) e non abbiano mai avuto altri codici che identificano malattie simili al morbo di Parkinson (332.1, 333.0, 333.1). L'altro si basa sulla frequenza di registrazione del codice 332.0 e sulla conferma del codice da parte di uno specialista. Ciascun algoritmo è stato testato usando la prescrizione di trattamento mirato al MP

TIPO DI STUDIO Coorte storico

## POPOLAZIONE IN STUDIO

Membri di una associazione di veterani che assicurava assistenza sanitaria. USA - Los Angeles

PERIODO IN STUDIO ottobre 2001 – settembre 2004

## DATABASE

Veterans Administration Greater Los Angeles Healthcare system

## DEFINIZIONE DI CASO

Tutti i soggetti che hanno avuto diagnosi di malattia di Parkinson (codice ICD-9 332.0) alla dimissione o ad un controllo ambulatoriale.

## STANDARD PER LA VALIDAZIONE

Conferma diagnostica di ciascun caso in base alla documentazione della cartella clinica, in base a cui i pazienti sono stati classificati in casi confermati come affetti dal morbo di P, possibili casi di MP e casi non confermati.

## MODALITÀ DI VALIDAZIONE

Confronto dei casi identificati in base ai codici ICD-9 su file amministrativi con i dati clinici.

## RISULTATI

PPV: aumenta da 75.6% (se il codice 332.0 è riportato 1 sola volta fino a 83.2% se il codice è riportato 5+ volte; il VPN si riduce.

Il VPP aumenta fino a 87.6% se quando il codice 332.0 è riportato più volte da uno specialista

neurologo e fino a 93.3% se riportato più volte da uno specialista dei disordini motori. Il NPV in questo caso rimane invariato

Il VPP per aver avuto una prescrizione farmaceutica specifica (qualunque farmaco) è di 85.5%. L'uso combinato con i farmaci porta il VPP dell'algoritmo di base all'88% e quello dell'algoritmo con diagnosi specialistica all'88,2%.

## CONCLUSIONI DEGLI AUTORI

L'uso del codice 332.0 nell'algoritmo che include le visite specialistiche con referti ripetuti nell'arco di 3 anni assieme alle prescrizioni di farmaci specifici, escludendo i pazienti con codici di MP mimetici, consente di ottenere i valori più elevati di VPP, cioè le probabilità più elevate di identificare i veri malati di MP.

#### COMMENTI RELATORE

L'obiettivo del lavoro non è direttamente pertinente al tema della validazione delle stime di occorrenza. Si direbbe più orientato alla identificazione nei database elettronici, di pazienti con elevata probabilità di essere un caso "vero" di MP. In questo contesto viene sviluppato un algoritmo integrato di più fonti che aumenta il VPP, cioè un algoritmo che identifica bene casi da arruolare in coorti di pazienti per una valutazione di esito o di performance assistenziale.

(146)

#### OBIETTIVI

Obiettivo principale : stimare la prevalenza della BPCO utilizzando i codici ICD-9 in un DB elettronico nazionale dei MMG.

Obiettivi secondari : distinguere la BPCO dall'asma; assegnare la gravità della BPCO (lieve, moderata, grave) mediante uso di farmaci.

#### POPOLAZIONE IN STUDIO

popolazione (N 5 milioni) assistita nel 1998 da 500 MMG del RU (UK) che partecipano ad una sorveglianza continua afferente al DB nazionale. Il campione è stato precedentemente stimato rappresentativo della popolazione generale, pur rappresentando il 10%.

#### PERIODO IN STUDIO

#### DATABASE

DB nazionale dei MMG

#### DEFINIZIONE DI CASO

i pazienti riportati nel DB dei MMG con almeno un codice del sistema di codifica nazionale associato alla BPCO, indipendentemente dalla presenza o meno di asma.

L'indice di severità include età, FEV<sub>1</sub>, numero di broncodilatatori e/o cortisonici prescritti in intervalli di 6 mesi.

#### MISURE DI VALIDAZIONE

k, sensibilità e specificità del DB usato rispetto alla diagnosi clinica del medico curante.

#### STANDARD PER LA VALIDAZIONE

La diagnosi e il grado di severità della BPCO attribuito dai MMG in base ai dati clinici ad un campione di 300 pazienti identificati nel DB come affetti da BPCO (incluse BPCO, enfisema, bronchite cronica) e/o asma.

#### MODALITÀ DI VALIDAZIONE

Stima dell'indice di concordanza K, sensibilità e specificità della diagnosi e del grado di severità della BPCO riportati nel file elettronico dei MMG confrontato con la diagnosi clinica del medico curante.

#### RISULTATI

La diagnosi di BPCO riportata nel file dei MMG ha  $k = 0.46$ ; sensibilità = 94,4%; specificità = 50%, quando confrontata con la diagnosi clinica del medico curante.

La attribuzione di severità della BPCO ha  $k = 0.46$  per le forme lievi;  $k = 0.59$  per le forme medie;  $k = 0,68$  per le forme gravi.

#### CONCLUSIONI DEGLI AUTORI

Gli algoritmi usati sul DB per identificare i pazienti con BPCO e assegnare la gravità della malattia danno stime valide, mentre la capacità di discriminare la BPCO dall'asma è accettabile solo per le forme medio-gravi della malattia.

#### COMMENTI RELATORE

Un buon esempio di validazione che usa un “gold standard clinico” e una definizione di caso completa. Il giudizio sui risultati di validazione è discutibile.

(147)

#### RAZIONALE

#### OBIETTIVI

Valutare la coerenza tra diverse regioni di 4 DB amministrativi rispetto per i dati su asma e BPCO e stimare la validità delle stime di occorrenza e di associazione con inquinanti ambientali

#### POPOLAZIONE IN STUDIO

Inghilterra (49 milioni di abitanti).

PERIODO IN STUDIO 1991-1995.

#### DATABASE

Registri nazionali di mortalità e ricoveri. Il DB nazionale dei MMG del 1994 (registro campionario che copre una popolazione di 5 milioni) 1994. I risultati campionari (20000 soggetti) dell'indagine sullo stato di salute della popolazione del 1995, relativi ai sintomi riferiti per asma o BPCO.

#### DEFINIZIONE DI CASO

Essere deceduto per una causa di morte registrata con codice ICD-9 493 o 490-492, 494-496;

Essere stato ricoverato d'urgenza con una diagnosi principale di dimissione registrata con codice ICD-9 493 o 490-492, 494-496;

L'aver ricevuto dal proprio medico curante, almeno una prescrizione per farmaci respiratori inalatori nell'ultimo anno e una diagnosi clinica di asma o BPCO nell'ultimo anno o precedentemente;

L'aver riferito sintomi quali fischi e sibili (per l'asma) o tosse e catarro per 3+ mesi all'anno (per la BPCO) durante l'intervista sullo stato di salute della popolazione.

#### MISURE DI OCCORRENZA

Tassi di prevalenza standardizzati per sesso ed età ottenuti dalle diverse fonti e stratificati per residenza (50 regioni e 4 aree di urbanizzazione: conurbana, urbana, mista, rurale).

#### STANDARD PER LA VALIDAZIONE

Non è stato definito uno standard di riferimento, le fonti di informazione si equivalgono.

#### MODALITÀ DI VALIDAZIONE

Correlazione tra le fonti per le stime di prevalenza di asma e BPCO per regione e tipologia delle aree di urbanizzazione.

#### RISULTATI

L'asma mostra valori poco coerenti tra le diverse fonti e una debole correlazione tra le regioni geografiche, mentre fonti diverse forniscono dati coerenti per la BPCO.

L'aggiustamento per stato sociale e fumo attenua debolmente le differenze geografiche osservate per la BPCO, invece non ha alcun effetto sul quadro dell'asma.

#### CONCLUSIONI DEGLI AUTORI

I dati per la BPCO data la coerenza tra le fonti di dati in ciascuna regione riflettono la prevalenza vera della malattia e possono essere utilizzati nelle analisi di impatto. Per l'asma, al contrario, la grande variabilità tra le fonti sconsiglia l'uso.

#### COMMENTI RELATORE

E' un buon esempio di utilizzo di analisi geografica, che risente dei limiti e dei vantaggi di questo

approccio, primo fra tutti non è facile distinguere le differenze di occorrenza della malattia dalle

differenze nella gestione dei database tra le aree.

Non si tratta di una vera e propria validazione in quanto non è stato definito e utilizzato uno standard con cui confrontare le fonti.

Le fonti sono confrontate in alternativa tra loro, mentre è sempre più frequente l'uso di DB integrati, in quanto osservano popolazioni di pazienti diverse.

I tre lavori che seguono (149, 150, 153) sono lavori comparsi su riviste di metodologia statistica, fra tanti altri, che qui riportiamo come esempio di approccio metodologico applicato soprattutto quando si intenda procedere a stime di accuratezza diagnostica (Sensibilità e Specificità) in assenza di gold standard.

Mentre è piuttosto semplice da applicare la procedura che si suggerisce per "aggiustare" la stima di prevalenza quando si disponga di un metodo di accertamento di nota inaccuratezza, un approccio esemplificato ed applicato praticamente in alcuni dei lavori già citati ad esempio per la stima di prevalenza di alcuni tipi di tumore, è più complessa la soluzione del problema, ovvero giungere ad una stima accurata di prevalenza quando, non disponendo di stime di accuratezza per mancanza di Gold Standard, si deve procedere basandosi su assunti più o meno plausibili nella realtà.

Una soluzione a questo problema è data nel caso si disponga di 3 procedure di accertamento fra loro condizionalmente indipendenti. In questo caso si dimostra come, potendo costruire 7 equazioni lineari indipendenti si possa procedere alla stima dei 7 parametri di interesse: Sensibilità e Specificità di ciascuno dei 3 metodi applicati, e stima di prevalenza di interesse. È questo l'approccio attraverso il modello di regressione a classi latenti applicato in alcuni dei lavori già considerati.

Qualora infine non si possa assumere l'indipendenza condizionale delle procedure di identificazione dei casi per stimare prevalenza e accuratezza, in assenza di gold standard, si può ricorrere, come trattato in (149) e (153), ad un approccio Bayesiano oppure, come descritto in (150), ad una analisi di classi latenti con modelli a effetti casuali.

## **CONCLUSIONE**

La revisione bibliografica ha consentito di evidenziare fondamentalmente due tipologie di riferimenti:

- 1) Riferimenti bibliografici orientati alla validazione dei criteri di estrazione dei casi dai database amministrativi per mezzo del confronto con un gold standard o quantomeno con altre fonti informative giudicate equivalenti dal punto di vista dell'accuratezza diagnostica. Nella maggior parte dei casi il confronto tra differenti fonti informative consente una stima dell'errore di classificazione; in alcuni casi il confronto consente una stima di un bias di selezione.
- 2) Riferimenti bibliografici che descrivono strumenti statistici particolari al fine di correggere l'errore di classificazione o il bias di selezione.

Facendo particolare riferimento alla letteratura di interesse clinico cardiologica possiamo segnalare, per quanto riguarda:

- 1) Validazione delle diagnosi. Un primo riferimento descrive il confronto tra le stime di incidenza per infarto miocardico acuto, angina instabile e scompenso cardiaco in Olanda, ottenute per mezzo del file dei ricoveri e del registro delle cause di morte, e le stime ottenute da un registro basato su una coorte di soggetti seguiti dal 1927. Quest'ultima fonte viene considerata il gold standard. I risultati vengono espressi in termini di sensibilità e VPP. Vengono riportati buoni valori di sensibilità e VPP per la cardiopatia ischemica in generale e per l'infarto miocardico acuto e valori meno

buoni per l'angina instabile e per lo scompenso <sup>(148)</sup>. Un secondo riferimento descrive il confronto tra il file dei ricoveri e le informazioni contenute nelle cartelle cliniche, che vengono considerate il gold standard per la diagnosi di infarto miocardico acuto e per l'angina instabile. I criteri diagnostici utilizzati per la valutazione delle cartelle cliniche sono quelli approvati dalla società scientifiche (AHA/ESC). Vengono riportati buoni valori di VPP del codice 410 per l'infarto e 411 per l'angina instabile <sup>(122)</sup>. La cartella clinica è stata utilizzata come gold standard anche per la validazione del codice di diagnosi di frattura di femore. In questo caso emerge una rilevante variabilità nell'accuratezza diagnostica del codice con percentuali di sottostima pari a 46% in un ospedale e di sovrastima del 19% in un altro ospedale, ai due estremi <sup>(28)</sup>. In altri riferimenti bibliografici non viene adottato un gold standard per la validazione diagnostica, ma viene fatto un confronto tra differenti fonti informative: per la stima di prevalenza di diabete, ipertensione arteriosa, asma, broncopneumopatia cronica ostruttiva <sup>(70,6)</sup> il confronto viene effettuato tra dati di ricovero, prestazioni ambulatoriali, registro della medicina di base (Primary healthcare), singolarmente presi o in combinazione; i dati vengono confrontati anche con alcuni riferimenti esterni al fine di valutare la plausibilità dei risultati ottenuti. Emerge inoltre la necessità di utilizzare informazioni su molti anni consecutivi al fine di ottenere stime di prevalenza stabili. Ulteriori confronti possono essere effettuati tra i database amministrativi e gli studi basati su questionari relativi allo stato di salute percepito. In uno studio <sup>(95)</sup> emerge una buona concordanza tra i due strumenti informativi per il diabete e l'ipertensione arteriosa; una concordanza più bassa per altre condizioni. In questo riferimento viene chiaramente enunciata la difficoltà a trovare un gold standard informativo per molte condizioni cliniche. In questo lavoro è utilizzata la statistica K di concordanza aggiusta per il caso per esprimere i risultati del confronto (misura di riproducibilità) anziché la stima della sensibilità o del VPP (misura di accuratezza). Risultati sovrapponibili vengono ottenuti confrontando fonti informative identiche in contesti differenti <sup>(2,145)</sup>. Altri tre riferimenti bibliografici consentono di approfondire alcuni aspetti complementari al problema della validazione. In un lavoro orientato ai criteri di estrazione dei casi di ICTUS cerebrali vengono confrontati algoritmi più o meno sensibili o specifici rispetto a un gold standard riconosciuto basato sulla lettura della cartella clinica. Emerge una ampia variabilità nel numero di casi estratti in funzione del tipo di algoritmo utilizzato, anche se il trend della stima di prevalenza appare stabile indipendentemente dal tipo di algoritmo utilizzato <sup>(102)</sup>. In un altro lavoro vengono confrontati i due sistemi di codifica ICD9-cm e ICD10 su 32 condizioni cliniche. Si evidenzia una sostanziale equivalenza complessiva della validità dei due sistemi di codifica, anche se ci possono essere differenze a vantaggio dell'uno o dell'altro sistema di codifica in relazione ad alcune condizioni cliniche; complessivamente l'introduzione del sistema di codifica ICD10 non sembra apportare un apprezzabile miglioramento nella validità dei codici <sup>(115)</sup>. Infine un richiamo all'esistenza del bias di selezione oltre agli errori di misclassificazione: utilizzando dati del registro delle cause di morte in aggiunta al file dei ricoveri e al registro di popolazione <sup>(151)</sup> si evidenzia che il 40% dei casi di infarto miocardico acuto in Olanda muoiono prima di raggiungere l'ospedale. Mentre l'errore di misclassificazione può essere opportunamente gestito con metodi di aggiustamento o per mezzo dell'analisi di sensibilità, il bias di selezione, se non riconosciuto, comporta una distorsione irrimediabile delle stime di incidenza e di prevalenza.

- 2) Metodi statistici. In realtà, sotto determinate condizioni (assunti di validità), è possibile correggere bias di selezione nella stima di prevalenza di malattie croniche

qualora si disponga di più fonti informative indipendenti. Ad esempio è possibile utilizzare metodi statistici che consentono di studiare la distribuzione temporale dei contatti con medici di Medicina Generale e metterla in relazione con la stima di prevalenza, stimando i contatti “mancati” a causa di una finestra temporale di studio troppo stretta utilizzando solo le informazioni “parziali” (distorte da selezione) ottenute dai contatti “avvenuti” <sup>(7)</sup> . Più frequenti sono i riferimenti a metodi statistici appropriati (sempre sotto determinati assunti di validità) a correggere bias di classificazione. Nella bibliografia selezionata risulta ad esempio un lavoro che applica a questo scopo metodi bayesiani (Bayesian latent class models) alla correzione dell’errore di classificazione nell’osteoartrite <sup>(84)</sup> e nella polimiosite/dermatomiosite <sup>(117)</sup> . L’approccio descritto prescinde dal ricorso ad un gold standard di riferimento. Sensibilità, specificità e prevalenza vengono infatti stimati come a parametri “latenti”, che non possono essere osservati direttamente, ma possono essere derivati dai dati disponibili combinando le informazioni a priori su tali parametri, modellate come distribuzioni di probabilità, con i dati campionari raccolti. Naturalmente la correttezza dei risultati che si ottengono dipende fortemente dalla validità degli assunti su cui il metodo si fonda. Non sempre né per ogni tipo di patologia né per ogni tipo di fonte informativa tale validità può essere dimostrata e, talora, nemmeno supposta.

- 3) Vale la pena infine di citare un riferimento bibliografico che descrive uno strumento (Check list) di valutazione della qualità dei database utilizzati come fonte informativa per gli studi osservazionali <sup>(11)</sup> .

## ALLEGATO E

### PATOLOGIA CONSIDERATA

1. Artrite, asma, diabete, malattie cardiache, ipertensione e ictus  
DB amministrativi vs dati campionari;  
2 metodi da valutare per concordanza e individuare determinanti
2. Lupus eritematoso sistemico: incidenza e prevalenza  
DB amministrativi; stime di sensibilità e specificità
3. Osteoporosi  
Algoritmo e metodo statistico che usa più fonti amministrative
4. Stroke  
Studio di fattibilità di sorveglianza sanitaria di una coorte con DB amministrativi
5. Frattura dell'anca nell'anziano  
Casistica identificata con SDO per monitorare complicanze e ricoveri
6. Diabete  
Programmare attività preventive e valutare esiti
7. Rassegna sistematica di problemi di stima con DB amministrativi
8. Diabete, ipertensione, cardiopatie ischemiche, ictus e ipercolesterolemia
9. Patologie croniche in genere  
DB di medicina generale e cross-validation
10. BPCO  
Validazione casi individuati con DB amministrativi
11. Diabete, ipertensione, BPCO e ulcera gastroduodenale  
Validazione casi individuati con indagine campionaria ISTAT vs DB SIMG (550 MMG)
12. Diabete, ipertensione, asma, BPCO  
Stime di prevalenza usando SDO e dati della medicina generale
13. Asma e BPCO combinate (MPCO)  
Stime di prevalenza e utilizzo dell'archivio prescrizioni
14. Diabete, ipertensione, asma, BPCO e patologie cerebrovascolari  
L'uso dei registri di attività sanitarie.
15. Osteoartrite  
Stime di prevalenza utilizzando diversi algoritmi e DB, modelli bayesiani (latent class models) e saggiando l'effetto di diverse assunzioni a priori.
16. Morbo di Parkinson  
Identificazione casi: con diversi algoritmi e DB (Medicare claims) vs self-reported (Medicare Current Beneficiary Survey).
17. Ipertensione  
Confronto di due algoritmi usando DB (rimborsi, piano assicurativo mid-Atlantic Health plan) vs cartelle cliniche
18. Ipertensione  
Stime di prevalenza usando diversi DB e algoritmi e validazione con differenti gold standard
19. BPCO  
Stime di prevalenza in DB nazionale MMG distinguendo asma e graduando la gravità in base alla prescrizione di farmaci
20. Asma e BPCO  
Confronto fra regioni utilizzando quattro DB e studiando la relazione con inquinamento ambientale.

## POPOLAZIONI

- 1 Manitoba (1.2 milioni)
- 2 Quebec (7.5 milioni)
- 3 Manitoba
- 4 Manitoba
- 5 Tre ospedali di Oslo
- 6 n.a.
- 7 n.a.
- 8 Manitoba: campione di 2.797 unità (survey condotta nel periodo 1989-90)
- 9 Paesi Bassi: 23.500 pazienti 1987-88 15 MMG Indagine Nazionale MG e malattie croniche
- 10 Quebec: 860.705 persone 65+ anni 1994-95
- 11 Italia: 2 campioni di popolazione di età > 14 anni 199.799 1999-2000 ISTAT; 432.747 DB-SIMG tra i pazienti vivi alla fine del 2000
- 12 Contea di Ostergotland (Svezia): 415.000 residenti 2003. Dati 1999-2003 prevalenza longitudinale (visti nei 5 anni precedenti e vivi a fine 2003)
- 13 Svezia: 6.756.453 >20 anni 2004
- 14 Municipio di Tierp (Svezia): 20.037 residenti 1996 registri sanitari
- 15 Quebec: >65 anni Regie de l'Assurance Maladie du Quebec nel 2002.
- 16 US: > 30.000 soggetti survey MCBS 1992-2000 archivi informatizzati dei rimborsi (claims) Medicare
- 17 US: popolazione inclusa piano assicurativo mid-Atlantic Health plan 1/1/2000 31/3/2003.
- 18 Ontario
- 19 UK: 5 milioni assistiti (10% popolazione nazionale) 1998 500 MMG sorveglianza continua DB nazionale.
- 20 UK: 49 milioni abitanti 1991-5 registro nazionale MMG (5 milioni) 1994. Campione di popolazione (20.000) 1995 sui sintomi riferiti.

## TIPO DI STUDIO

1. Trasversale
2. Trasversale
3. Coorte retrospettivo
4. Metanalisi
5. Survey su registrazioni di un anno
6. Rassegna bibliografica
7. Rassegna metodologica
8. Prevalenza in survey vs DB
9. Coorte prospettico
10. Prevalenza DB vs "gold standard"
11. Studio descrittivo prevalenza da 2 DB e studio esplorativo associazione età prevalenza
12. Studio descrittivo prevalenza stimata con algoritmi su 3 DB
13. Stima di incidenza patologie respiratorie croniche
14. Stima di prevalenza
15. Studio di casistica
16. Studio di casistica
17. Studio di casistica
18. Studio di prevalenza con linkage DB diversi anni
19. Studio di incidenza BPCO con validazione clinica
20. Validazione fonti dati

## DATABASE UTILIZZATI

1. Dimissioni ospedaliere, rimborsi medici, prescrizioni farmaci: 152 definizioni di casi da 1 o più fonti (algoritmi in precedenti lavori)
2. Dimissioni ospedaliere MEDECHO I (almeno una diagnosi 710.0), file demografico RAMQ, file rimborsi medici RAMQ (codice 710.0), linkati con id sanitario unico (incidenza 1994, prevalenza 2003).
3. Dimissioni ospedaliere, rimborsi medici per osteoporosi (733) e fratture (820-21; 805; 812; 813-14), farmaceutica, test BMI, prestazioni ambulatoriali (ADG) e corticosteroidi per comorbidità
4. Research Data Repository con SDO, rimborsi medici, prescrizioni farmaceutiche
5. Ufficio Centrale ricoveri ospedalieri, cartelle cliniche e registri sale operatorie
6. Ricerca di letteratura in Medline
7. n.a.
8. Rispondenti survey su patologie cardio- e cerebro-vascolari
9. Survey pazienti 15 medici: diabete, ipertensione, ischemie del miocardio, asma, BPCO, osteoartrite ginocchio e anca.
10. Rimborsi prestazioni mediche e farmaceutiche
11. VI indagine campionaria ISTAT e database SIMG
12. Dimissioni ospedaliere, prestazioni ambulatoriali, dati MMG
13. Due database farmaceutica: uno con dati aggregati per età e sesso, tipo e dosi giornaliere del farmaco, altro nominativo con le stesse informazioni: confronto entro contea
14. Registri MMG, Medicina del Lavoro, ricoveri e visite in ambulatori ospedalieri
15. Regie de l'Assurance Malarie du Quebec, rimborsi farmaceutica 3 algoritmi:
  - 1 almeno una diagnosi clinica di Osteoartrite ;
  - 2 almeno una prescrizione per OA;
  - 3 almeno una procedura medico chirurgica
16. Database rimborsi Medicare 1992-2000 4 algoritmi:
  - ICD9=332.0;
  - ICD9 332.0, 332.1, 333.0, 333.1
  - solo sui rimborsi del medico curante
  - inclusi rimborsi ad altri medici
17. Database rimborsi assicurativi 2000-2003 2 algoritmi:
  - 1 rimborsi medici (ICD9 HTN -401.xx);
  - 2 rimborsi medici e prescrizioni farmaceutiche.Casi incidenti: escludendo medico o farmacologico 12 mesi precedenti
18. Dimissioni ospedaliere, rimborsi medici. Combinazioni codici dai due archivi vs dati campione medici usando come gold standard DB elettronico nazionale MMG
19. Mortalità e ricoveri urgenti (ICD-9 493 asma; 490-492, 494,496 BPCO), diagnosi MMG, sintomi riferiti indagine stato di salute (fischi e sibili per asma, tosse e catarro per 3+ mesi all'anno per COPD)

## MODALITÀ DI VALIDAZIONE E GOLD STANDARD

1. CCHS stime campionarie (campionamento stratificato multi-stadio): 8.120 casi nel 2001
2. Validazione interna confronto stime prodotte dalle fonti separatamente
3. Test BMD diagnostico richiesto dai medici (T score  $\leq 2.5$  in donna adulta)
4. Confronto dati amministrativi con dati rimborsi medici. Analisi di sensibilità
5. SDO frattura anca vs cartelle cliniche e registri sale operatorie. Verifica manuale medici
6. Stime di sensibilità, specificità, VPP, stima della statistica kappa
7. n.a.
8. G.S. diagnosi DB rimborsi ospedali e medici nei tre anni precedenti l'intervista
9. Stime di prevalenza basate su periodo di osservazione di 600 gg dello stesso database dei 15 MMG aderenti allo studio
10. G.S. prevalenza indagine sullo stato di salute 1994/1995, analisi di sensibilità con/senza codice 490. Casi possibili vs probabili.
11. Stima di occorrenza da DB SIMG
12. Stime da studi precedenti: WHO asma; INTERGENE ipertensione; OLIN asma e BPCO
13. Precedenti survey con prevalenza 4-6% per BPCO e 8% asma.
14. Nessuno
15. Prevalenza (IC 95%) aggiustata per misclassificazione (latent class models) sulla base di informazioni a priori (interviste a medici)
16. A) pazienti survey MCBS PD autoriferita  
B): + prescrizione per farmaci PD
17. Cartelle cliniche MMG, internisti, cardiologi, nefrologi: campione medici reclutati pazienti (cluster)
18. 1) campione di medici di base, da cui campione casuale di pazienti 35-64  
2) Survey 2001 diagnosi autoriferita
19. Diagnosi e gravità BPCO confermata MMG su campione pazienti identificati nel DB affetti BPCO e/o asma (enfisema, bronchite cronica) indipendentemente dai criteri diagnostici
20. Coerenza tra le fonti dei dati per le due malattie nelle diverse aree e regioni

## PROCEDURE DI VALIDAZIONE

1. Linkage deterministico tra archivi amministrativi e survey attraverso id sanitario unico anonimo. 83.9% linkati; statistica K
2. Sensibilità e specificità approccio Bayesiano gerarchico senza gold standard. 3 metodi per ricerca casi: 1 rimborsi medici, 2 dimissioni ospedaliere, 3 procedure / trattamenti per SLE
3. Training cohort 2000-01 donne 50+ con almeno un test BMD selezionati casi e non casi (T score) e validazione 1999-2000. 3 approcci: reti neurali, classificazione ad albero, regressione logistica.
4. Controllo manuale delle diagnosi considerando le cartelle e i registri di chirurgia come gold standard
5. Rilevanza, validità e utilità del database criteri di eleggibilità DB: presentati i metodi di valutazione
6. N.A.
7. N.A.
8. Concordezza fonti, sensibilità, specificità, valore predittivo positivo e negativo survey vs rimborsi
9. Validazione cross-sectional modelli con
  - 1) data primo contatto
  - 2) totale pazienti e contatti,
  - 3) pazienti osservati in diversi intervalli tempo
10. Confronto prevalenze con i due approcci: DB vs Indagine di Popolazione; analisi di sensibilità codici ICD-9 BPCO con o senza il 490 su BPCO, probabili possibili, e non rispondenti a definizione di caso.
11. Regressione lineare tra età e differenze dell'occorrenza
12. Confronto con studi precedenti e analisi di sensibilità per diversi algoritmi di selezione e diversi DB
13. Confronto survey precedenti Svezia. Coefficienti di correlazione tra i due registri per i farmaci respiratori selezionati.
14. Analisi di sensibilità delle stime annuali vs. stime longitudinali su due e tre anni modelli bayesiani
15. 4 algoritmi di cui si valutano sensibilità, specificità, PPV rispetto a ciascun reference standard e stima costi annuali per Parkinsonismo.
16. 2 algoritmi: sensibilità, specificità, statistica kappa
17. Confronto 12 algoritmi con dati MG selezionati e survey: K, specificità, sensibilità, VPP, VPN
18. K, sensibilità e specificità della diagnosi e del grado di severità della BPCO
19. Correlazione di Spearman tra i tassi standardizzati sesso/età ottenuti dalle varie fonti e stratificati per residenza (50 regioni e 4 aree di urbanizzazione: conurbana, urbana, mista, rurale)

## BIBLIOGRAFIA

- 1 Rawson NS, Malcolm E, D'Arcy C  
Reliability of the recording of schizophrenia and depressive disorder in the Saskatchewan health care datafiles  
Soc Psychiatry Psychiatr Epidemiol. 1997 May;32(4):191-9
- 2 Robinson JR, Young TK, Roos LL, Gelskey DE  
Estimating the burden of disease. Comparing administrative data and self-reports  
Med Care. 1997 Sep;35(9):932-47**
- 3 Rawson NS, D'Arcy C  
Assessing the validity of diagnostic information in administrative health care utilization data: experience in Saskatchewan  
Pharmacoepidemiol Drug Saf. 1998 Nov;7(6):389-98
- 4 West SL, Richter A, Melfi CA, McNutt M, Nennstiel ME, Mauskopf JA  
Assessing the Saskatchewan database for outcomes research studies of depression and its treatment  
J Clin Epidemiol. 2000 Aug;53(8):823-31
- 5 Wigertz A, Westerling R  
Measures of prevalence: which healthcare registers are applicable?  
Scand J Public Health. 2001 Mar;29(1):55-62**
- 6 Wong O.  
Use of administrative databases in occupational and environmental epidemiologic research  
Ann Epidemiol. 2001 Jul;11(5):281-5**
- 7 Hoogenveen R, Westert G, Dijkgraaf M, Schellevis F, de Bakker D  
Disease prevalence estimations based on contact registrations in general practice  
Stat Med. 2002 Aug 15;21(15):2271-85**
- 8 Rodrigues EJ, Simpson E, Richard H, Pilote L  
Regional variation in the management of acute myocardial infarction in the province of Quebec  
Can J Cardiol. 2002 Oct;18(10):1067-76
- 9 Finnerty M, Altmansberger R, Bopp J, Carpinello S, Docherty JP, Fisher W, Jensen P, Krishnan P, Mittleman M, Olfson M, Tricarico J, White T, Felton C  
Using state administrative and pharmacy data bases to develop a clinical decision support tool for schizophrenia guidelines.  
Schizophr Bull. 2002;28(1):85-94
- 10 Mike Thomas, Jennifer Cleland and David Prince  
Database studies in asthma pharmacoconomics: uses, limitations and quality makers  
Expert Opin. Pharmacother. (2003) 4(3): 351-358
- 11 Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, Martin BC, Stang P.  
A checklist for retrospective database studies--report of the ISPOR Task Force on Retrospective Databases  
Value Health. 2003 Mar-Apr;6(2):90-7**

- 12**  
**AP**     **Cricelli C, Mazzaglia G, Samani F, Marchi M, Sabatini A, Nardi R, Ventriglia G, Caputi**  
**Prevalence estimates for chronic diseases in Italy: exploring the differences between self-report and primary care databases**  
**J Public Health Med. 2003 Sep;25(3):254-8**
- 13     Jaffe AB, Levine J.  
Efficacy and effectiveness of first- and second-generation antipsychotics in schizophrenia  
J Clin Psychiatry. 2003;64 Suppl 17:3-6
- 14     Monesi L, Fernandez CR, D'Ettorre A, Romero M, Sasso E, Lepore V, Tognoni G, Monesi G, Ferrarese A.  
Administrative data as source for epidemiological research: clinical pathways of diabetic patients  
Assist Inferm Ric. 2003 Apr-Jun;22(2):81-90
- 15     White RH, Zhou H, Romano PS  
Incidence of symptomatic venous thromboembolism after different elective or urgent surgical procedures  
Thromb Haemost. 2003 Sep;90(3):446-55
- 16     Kearon C.  
Epidemiology of postoperative venous thromboembolism: lessons from an administrative data base  
Thromb Haemost. 2003 Sep;90(3):367-8
- 17     Goldner EM, Jones W, Waraich P.  
Using administrative data to analyze the prevalence and distribution of schizophrenic disorders  
Psychiatr Serv. 2003 Jul;54(7):1017-21
- 18**     **Saydah SH, Geiss LS, Tierney E, Benjamin SM, Engelgau M, and Brancati F**  
**Review of the Performance of Methods to Identify Diabetes Cases among Vital Statistics, Administrative, and Survey Data**  
**Ann Epidemiol 2004;14:507–516.**
- 19     Peabody JW, Luck J, Jain S, Bertenthal D, Glassman P  
Assessing the accuracy of administrative data in health information systems  
Med Care. 2004 Nov;42(11):1066-72
- 20     Whitehorn D, Richard JC, Kopala LC  
Hospitalization in the first year of treatment for schizophrenia  
Can J Psychiatry. 2004 Sep;49(9):635-8
- 21     Steele LS, Glazier RH, Lin E, Evans M  
Using administrative data to measure ambulatory mental health service provision in primary care  
Med Care. 2004 Oct;42(10):960-5
- 22     Levy AR, Tamblyn RM, Abrahamowicz M, McLeod PJ, Fitchett D  
Use of time-dependent measures to estimate benefits of beta-blockers after myocardial infarction  
Pharmacoepidemiol Drug Saf. 2004 Sep;13(9):623-31
- 23     Herbert MA, Prince SL, Williams JL, Magee MJ, Mack MJ  
Are unaudited records from an outcomes registry database accurate?

- Ann Thorac Surg. 2004 Jun;77(6):1960-4; discussion 1964-5
- 24 Thomas C, Cadwallader HL, Riley TV  
Surgical-site infections after orthopaedic surgery: statewide surveillance using linked administrative databases.  
J Hosp Infect. 2004 May;57(1):25-30
- 25 Quan H, Parsons GA, Ghali WA.  
Assessing accuracy of diagnosis-type indicators for flagging complications in administrative data  
J Clin Epidemiol. 2004 Apr;57(4):366-72.
- 26 Reilly JJ, Chin B, Berkowitz J, Weedon J, Avitable M  
Use of a state-wide administrative database in assessing a regional trauma system: the New York City experience.  
J Am Coll Surg. 2004 Apr;198(4):509-18
- 27 Schneeweiss S, Avorn J  
A review of uses of health care utilization databases for epidemiologic research on therapeutics  
JCE, 2005, 58: 323 – 337
- 28 Lofthus CM, Cappelenb I, Osnes EK, Falcha JA, Kristiansend IS, Medhuse AW, Nordslettenc L, Meyerb HE**  
**Local and national electronic databases in Norway demonstrate a varying degree of validity**  
**JCE, 2005, 58: 280 – 285**
- 29 J Raftery, P Roderick and A Stevens  
Potential use of routine databases in health technology assessment  
Health Technology Assessment, 2005, 9: No. 20
- 30 Wunsch H, Harrison DA, Rowan K  
Health services research in critical care using administrative data  
J Crit Care. 2005 Sep;20(3):264-9.
- 31 Leake JL, Werneck RI  
The use of administrative databases to assess oral health care  
J Public Health Dent. 2005 Winter;65(1):21-35
- 32 Byrne N, Regan C, Howard L  
Administrative registers in psychiatric research: a systematic review of validity studies  
Acta Psychiatr Scand. 2005 Dec;112(6):409-14
- 33 Lacasse Y, Montori VM, Lanthier C, Maltis F**  
**The validity of diagnosing chronic obstructive pulmonary disease from a large administrative database**  
**Can Respir J. 2005 Jul-Aug;12(5):251-6**
- 34 Sikka R, Xia F, Aubert RE  
Estimating medication persistency using administrative claims data  
Am J Manag Care. 2005 Jul;11(7):449-57
- 35 Aronsky D, Haug PJ, Lagor C, Dean NC  
Accuracy of administrative data for identifying patients with pneumonia  
Am J Med Qual. 2005 Nov-Dec;20(6):319-28

- 36 Blackburn DF, Dobson RT, Blackburn JL, Wilson TW  
Cardiovascular morbidity associated with nonadherence to statin therapy  
Pharmacotherapy. 2005 Aug;25(8):1035-43
- 37 Beck CA, Richard H, Tu JV, Pilote L  
Administrative Data Feedback for Effective Cardiac Treatment: AFFECT, a cluster randomized trial  
JAMA. 2005 Jul 20;294(3):309-17
- 38 Mack MJ, Herbert M, Prince S, Dewey TM, Magee MJ, Edgerton JR.  
Does reporting of coronary artery bypass grafting from administrative databases accurately reflect actual clinical outcomes?  
J Thorac Cardiovasc Surg. 2005 Jun;129(6):1309-17
- 39 Blackburn DF, Dobson RT, Blackburn JL, Wilson TW, Stang MR, Semchuk WM  
Adherence to statins, beta-blockers and angiotensin-converting enzyme inhibitors following a first cardiovascular event: a retrospective cohort study  
Can J Cardiol. 2005 May 1;21(6):485-8
- 40 Percudani M, Barbui C, Fortino I, Petrovich L.  
Epidemiology of first- and second-generation antipsychotic agents in Lombardy, Italy.  
Pharmacopsychiatry. 2005 May;38(3):128-31
- 41 Raftery J, Roderick P, Stevens A**  
**Potential use of routine databases in health technology assessment.**  
**Health Technol Assess. 2005 May;9(20):1-92, iii-iv**
- 42 Humphries KH, Carere RG, Izadnegahdar M, Galbraith PD, Knudtson ML, Ghali WA**  
**Cross-provincial use of cardiac services: the importance of data-sharing for clinical registries and outcomes research**  
**Can J Cardiol. 2005 Mar;21(3):267-72**
- 43 Johnson D  
The utility of administrative data in helping the clinician understand and treat community-acquired pneumonia  
Respir Care Clin N Am. 2005 Mar;11(1):67-86
- 44 Ward MM.  
Estimating rare disease prevalence from administrative hospitalization databases.  
Epidemiology. 2005 Mar;16(2):270-1
- 45 De Coster C, Quan H, Finlayson A, Gao M, Halfon P, Humphries KH, Johansen H, Lix LM, Luthi JC, Ma J, Romano PS, Roos L, Sundararajan V, Tu JV, Webster G, Ghali WA  
Identifying priorities in methodological research using ICD-9-CM and ICD-10 administrative data: report from an international consortium  
BMC Health Services Research 2006, 1 - 6
- 46 Hennessy S  
Use of Health Care Databases in Pharmacoepidemiology  
Basic & Clinical Pharmacology & Toxicology, 2006, 98: 311–313
- 47 Lix L, Yogendran M, Burchill C, Metge C, McKeen N, Moore D, Bond R  
Defining and Validating Chronic Diseases: An Administrative Data Approach  
Manitoba Centre for Health Policy 2006 Department of Community Health Sciences

- 48 Vera Etches, John Frank, Erica Di Ruggiero, Doug Manuel  
MEASURING POPULATION HEALTH: A Review of Indicators  
Annu. Rev. Public Health 2006. 27:29–55
- 49 Fortuna D, Vizioli M, Contini A, Ghidoni I, Biagi B, Pacini D, Carnelos G, Grilli R  
Assessing clinical performance in cardiac surgery. Does a specialised clinical database make a difference?  
Interact Cardiovasc Thorac Surg. 2006 Apr;5(2):123-7
- 50 Hess LM, Raebel MA, Conner DA, Malone DC  
Measurement of adherence in pharmacy administrative databases: a proposal for standard definitions and preferred measures  
Ann Pharmacother. 2006 Jul-Aug;40(7-8):1280-88
- 51 Andrade SE, Kahler KH, Frech F, Chan KA  
Methods for evaluation of medication adherence and persistence using automated databases  
Pharmacoepidemiol Drug Saf. 2006 Aug;15(8):565-74; discussion 575-7
- 52 Grymonpre R, Cheang M, Fraser M, Metge C, Sitar DS  
Validity of a prescription claims database to estimate medication adherence in older persons  
Med Care. 2006 May;44(5):471-7
- 53 Gazmararian JA, Kripalani S, Miller MJ, Echt KV, Ren J, Rask K  
Factors associated with medication refill adherence in cardiovascular-related diseases: a focus on health literacy  
J Gen Intern Med. 2006 Dec;21(12):1215-21
- 54 Weiner MG, Garvin JH, Ten Have TR  
Assessing the accuracy of diagnostic codes in administrative databases: the impact of the sampling frame on sensitivity and specificity  
AMIA Annu Symp Proc. 2006:1140
- 55 Tromp M, Reitsma JB, Ravelli AC, Méray N, Bonsel GJ  
Record linkage: making the most out of errors in linking variables  
MIA Annu Symp Proc. 2006:779-83
- 56 Sousa P, Bazeley M, Johansson S, Wijk H  
The use of national registries data in three European countries in order to improve health care quality  
Int J Health Care Qual Assur Inc Leadersh Health Serv. 2006;19(6-7):551-60
- 57 Goldberg M  
Administrative data bases: could they be useful for epidemiology?  
Rev Epidemiol Sante Publique. 2006 Sep;54(4):297-303
- 58 Couris CM, Gutknecht C, Ecochard R, Gelas-Dore B, Hajri T, Colin C, Schott AM  
Estimates of the number of cancer patients hospitalized in a geographic area using claims data without a unique personal identifier  
Methods Inf Med. 2006;45(5):515-22
- 59 Wu EQ, Shi L, Birnbaum H, Hudson T, Kessler R  
Annual prevalence of diagnosed schizophrenia in the USA: a claims data analysis approach  
Psychol Med. 2006 Nov;36(11):1535-40.

- 60 Van Brabandt H, Camberlin C, Vrijens F, Parmentier Y, Ramaekers D, Bonneux L  
More is not better in the early care of acute myocardial infarction: a prospective cohort analysis on administrative databases  
Eur Heart J. 2006 Nov;27(22):2649-54
- 61 Bullano MF, Kamat S, Willey VJ, Barlas S, Watson DJ, Brenneman SK  
Agreement between administrative claims and the medical record in identifying patients with a diagnosis of hypertension  
Med Care. 2006 May;44(5):486-90**
- 62 Monte S, Fanizza C, Romero M, Rossi E, De Rosa M, Tognoni G  
Administrative databases as a basic tool for the epidemiology of cardiovascular diseases  
G Ital Cardiol (Rome). 2006 Mar;7(3):206-16
- 63 Guller U  
Surgical outcomes research based on administrative data: inferior or complementary to prospective randomized clinical trials?  
World J Surg. 2006 Mar;30(3):255-66
- 64 Mirandola M, Andretta M, Corbari L, Sorio A, Nosè M, Barbui C  
Prevalence, incidence and persistence of antipsychotic drug prescribing in the Italian general population: retrospective database analysis, 1999-2002.  
Pharmacoepidemiol Drug Saf. 2006 Jun;15(6):412-20
- 65 Gold HT, Do HT  
Evaluation of Three Algorithms to Identify Incident Breast Cancer in Medicare Claims Data  
Health Research and Educational Trust, 2007, 42: 2056 – 2069
- 66 Terris DD, Litaker DG, Koroukian SM  
Health state information derived from secondary databases is affected by multiple sources of bias  
JCE, 2007, 60: 734 - 741
- 67 Bernatsky S, Joseph L, Pineau CA, Tamblyn R, Feldman DE, Clarke AE  
A population-based assessment of systemic lupus erythematosus incidence and prevalence results and implications of using administrative data for epidemiological studies  
Rheumatology, 2007, 46: 1814 – 1818**
- 68 Madsen M, Gudnason V, Paja A, Palmieri L, Rocha E, Salomaa V, Sans S, Steinbach K, Vanuzzo D. Population-based register of acute myocardial infarction: manual of operations  
European Journal of Cardiovascular Prevention and Rehabilitation 2007, 14:S3–S22
- 69 Simona Giampaolia, Niklas Hammarb, Roza Adanyc and Christine De Perettid  
Population-based register of stroke: manual of operations  
European Journal of Cardiovascular Prevention and Rehabilitation 2007, 14 (Suppl 3):S23–S41
- 70 Wiréhn AB, Karlsson HM, Carstensen JM  
Estimating disease prevalence using a population-based administrative healthcare database  
Scand J Public Health. 2007;35(4):424-31**
- 71 Khanna R, Smith MJ

- Utilization and costs of medical services and prescription medications for rheumatoid arthritis among recipients covered by a state Medicaid program: a retrospective, cross-sectional, descriptive, database analysis  
Clin Ther. 2007 Nov;29(11):2456-67
- 72 Suissa S, Garbe E  
Primer: administrative health databases in observational studies of drug effects--advantages and disadvantages  
Nat Clin Pract Rheumatol. 2007 Dec;3(12):725-32
- 73 Ionescu-Iltu R, McCusker J, Ciampi A, Vadeboncoeur AM, Roberge D, Larouche D, Verdon J, Pineault R.  
Continuity of primary care and emergency department utilization among elderly people  
CMAJ. 2007 Nov 20;177(11):1362-8
- 74 Gini R, Capon A, Roti L, Mastromattei A, Buiatti E  
Femur fractures among elderly in Lazio and Tuscany regions from 1999 to 2003  
Epidemiol Prev. 2007 Jul-Aug;31(4):197-203
- 75 Welke KF, Peterson ED, Vaughan-Sarrazin MS, O'Brien SM, Rosenthal GE, Shook GJ, Dokholyan RS, Haan CK, Ferguson TB Jr.  
Comparison of cardiac surgery volumes and mortality rates between the Society of Thoracic Surgeons and Medicare databases from 1993 through 2001  
Ann Thorac Surg. 2007 Nov;84(5):1538-46
- 76 Nielsen KM, Foldspang A, Larsen ML, Gerdes LU, Rasmussen S, Faergeman O  
Estimating the incidence of the acute coronary syndrome: data from a Danish cohort of 138  
290 persons  
Eur J Cardiovasc Prev Rehabil. 2007 Oct;14(5):608-14
- 77 Myers RP, Leung Y, Shaheen AA, Li B.  
Validation of ICD-9-CM/ICD-10 coding algorithms for the identification of patients with acetaminophen overdose and hepatotoxicity using administrative data  
BMC Health Serv Res. 2007 Oct 2;7:159
- 78 Gold HT, Do HT  
Evaluation of three algorithms to identify incident breast cancer in Medicare claims data  
Health Serv Res. 2007 Oct;42(5):2056-69
- 79 Hyman N, Healey C, Osler T, Cataldo P  
Understanding variation in the management of rectal cancer: the potential of a surgeon-initiated database  
Am J Surg. 2007 Oct;194(4):559-62
- 80 Haukka J, Suvisaari J, Tuulio-Henriksson A, Lönnqvist J  
High concordance between self-reported medication and official prescription database information.  
Eur J Clin Pharmacol. 2007 Nov;63(11):1069-74
- 81 Shah BR, Hux JE, Laupacis A, Zinman B, Cauch-Dudek K, Booth GL  
Administrative data algorithms can describe ambulatory physician utilization.  
Health Serv Res. 2007 Aug;42(4):1783-96
- 82 Mohammed MA, Stevens A  
The value of administrative databases.  
BMJ. 2007 May 19;334(7602):1014-5

- 83 Aylin P, Bottle A, Majeed A  
Use of administrative data or clinical databases as predictors of risk of death in hospital: comparison of models  
BMJ. 2007 May 19;334(7602):1044.
- 84 Ladouceur M, Rahme E, Pineau CA, Joseph L**  
**Robustness of prevalence estimates derived from misclassified data from administrative databases**  
**Biometrics. 2007 Mar;63(1):272-9**
- 85 Roberts RJ, Stockwell DC, Slonim AD  
Discrepancies in administrative databases: implications for practice and research  
Crit Care Med. 2007 Mar;35(3):949-50
- 86 Shahian DM, Silverstein T, Lovett AF, Wolf RE, Normand SL  
Comparison of clinical and administrative data sources for hospital coronary artery bypass graft surgery report cards.  
Circulation. 2007 Mar 27;115(12):1518-27.
- 87 Bryson CL, Au DH, Young B, McDonnell MB, Fihn SD  
A refill adherence algorithm for multiple short intervals to estimate refill compliance (ReComp).  
Med Care. 2007 Jun;45(6):497-504
- 88 ten Berg MJ, Huisman A, van den Bemt PM, Schobben AF, Egberts AC, van Solinge WW.  
Linking laboratory and medication data: new opportunities for pharmacoepidemiological research  
Clin Chem Lab Med. 2007;45(1):13-9
- 89 Noyes K, Liu H, Holloway R, Dick AW**  
**Accuracy of Medicare claims data in identifying Parkinsonism cases: comparison with the Medicare current beneficiary survey.**  
**Mov Disord. 2007 Mar 15;22(4):509-14.**
- 90 Karmel R, Gibson D  
Event-based record linkage in health and aged care services data: a methodological innovation  
BMC Health Serv Res. 2007 Sep 25;7:154
- 91 Tu K, Chen Z, Lipscombe LL  
Prevalence and incidence of hypertension from 1995 to 2005: a population-based study  
CMAJ, 2008, 178: 1429 – 1435
- 92 Tu K, Chen Z, Lipscombe LL  
Mortality among patients with hypertension from 1995 to 2005: a population-based study  
CMAJ, 2008, 178: 1436 – 1440
- 93 Leenen FH, Dumais J, McInnis NH, Turton P, Stratyckuk L, Nemeth K, Moy Lum-Kwong M, Fodor G  
Results of the Ontario Survey on the Prevalence and Control of Hypertension  
CMAJ, 2008, 178: 1441 – 1449
- 94 Mohan S, Campbell NRC  
Hypertension management in Canada: good news, but important challenges remain  
CMAJ, 2008, 178: 1458 – 1460

- 95 Lix LM, Yogendran MS, Shaw SY, Burchill C, Metge C, Bond R  
Population-based data sources for chronic disease surveillance  
Chronic Diseases in Canada, 2008, 29: 31 – 38**
- 96 Leal J, Laupland KB  
Validity of electronic surveillance systems: a systematic review  
Journal of Hospital Infection, 2008, 69: 220 – 229
- 97 Couris CM, Polazzi S, Olive F, Remontet L, Bossard N, Gomez F, Schott AM, Mitton N, Colonna M, Trombert B  
Breast cancer incidence using administrative data: correction with sensitivity and specificity  
JCE, 2008 – In press
- 98 Patkar NM, Jeffrey R, Curtis JR, Teng GG, Allison JJ, Saag M, Martin C, Saag KG  
Administrative codes combined with medical records based criteria accurately identified bacterial infections among rheumatoid arthritis patients  
JCE, 2008 – In press
- 99 Baldi I, Vicari P, Di Cuonzo D, Zanetti R, Pagano E, Rosato R, Sacerdote C, Segnan N, Merletti F, Ciccone G  
A high positive predictive value algorithm using hospital administrative data identified incident cancer cases  
JCE, 2008, 61: 373 - 379
- 100 Lix LM, Yogendran MS, Leslie WD, Shaw SY, Baumgartner R, Bowman C, Metge C, Gumel A, Hux J, James RC  
Using multiple data features improved the validity of osteoporosis case ascertainment from administrative databases  
JCE, 2008, 61: 1250 – 1260**
- 101 Tricco AC, Pham B, Rawson NSB  
Manitoba and Saskatchewan administrative health care utilization databases are used differently to answer epidemiologic research questions  
JCE, 2008, 61: 192 – 197
- 102 Moore DF, Lix LM, Yogendran, MS, Martens P, Tamayo A  
Stroke surveillance in Manitoba, Canada: Estimates from administrative databases  
Chronic Diseases in Canada, 2008, 29: 22 – 30**
- 103 Welke KF, Karamlou T, Diggs BS  
Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease--a comparison of administrative and clinical data  
Cardiol Young. 2008 Dec;18 Suppl 2:137-44
- 104 LaRovere JM, Jeffries HE, Sachdeva RC, Rice TB, Wetzel RC, Cooper DS, Bird GL, Ghanayem NS, Checchia PA, Chang AC, Wessel DL  
Databases for assessing the outcomes of the treatment of patients with congenital and paediatric cardiac disease--the perspective of critical care  
Cardiol Young. 2008 Dec;18 Suppl 2:130-6
- 105 Strickland MJ, Riehle-Colarusso TJ, Jacobs JP, Reller MD, Mahle WT, Botto LD, Tolbert PE, Jacobs ML, Lacour-Gayet FG, Tchervenkov CI, Mavroudis C, Correa A  
The importance of nomenclature for congenital cardiac disease: implications for research and evaluation  
Cardiol Young. 2008 Dec;18 Suppl 2:92-100

- 106 Miller ML, Wang MC  
Accuracy of ICD-9-CM coding of cervical spine fractures: Implications for research using administrative databases  
Annu Proc Assoc Adv Automot Med. 2008;52:101-5
- 107 Andersohn F, Garbe E  
Pharmacoepidemiological research with large health databases  
Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008 51:1135-44
- 108 Hoffmann F, Andersohn F, Giersiepen K, Scharnetzky E, Garbe E  
Validation of secondary data. Strengths and limitations  
Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008 51:1118-26
- 109 Schubert I, Köster I, Küpper-Nybelen J, Ihle P  
Health services research based on routine data generated by the SHI. Potential uses of health insurance fund data in health services research  
Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz. 2008 51:1095-105
- 110 Hawker G, Guan J, Judge A, Dieppe P  
Knee arthroscopy in England and Ontario: patterns of use, changes over time, and relationship to total knee replacement  
J Bone Joint Surg Am. 2008 Nov;90(11):2337-45
- 111 Austin PC, Tu JV, Ko DT, Alter DA  
Factors associated with the use of evidence-based therapies after discharge among elderly patients with myocardial infarction.  
CMAJ. 2008 Oct 21;179(9):901-8
- 112 Austin PC, Tu JV, Ko DT, Alter DA  
Use of evidence-based therapies after discharge among elderly patients with acute myocardial infarction  
CMAJ. 2008 Oct 21;179(9):895-900
- 113 Rosato S, D'Errigo P, Badoni G, Fusco D, Perucci CA, Seccareccia F  
Comparison between administrative and clinical databases in the evaluation of cardiac surgery performance  
G Ital Cardiol (Rome). 2008 Aug;9(8):569-78
- 114 Austin PC, Donovan L, Yun L, Tu JV  
Comparing clinical and administrative data for profiling hospitals on postdischarge medication use by patients with acute myocardial infarction  
Am Heart J. 2008 Sep;156(3):595-605.
- 115 Quan H, Li B, Saunders LD, Parsons GA, Nilsson CI, Alibhai A, Ghali WA  
Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database  
Health Serv Res. 2008 Aug;43(4):1424-41**
- 116 Remontet L, Mitton N, Couris CM, Iwaz J, Gomez F, Olive F, Polazzi S, Schott AM, Trombert B, Bossard N, Colonna M  
Is it possible to estimate the incidence of breast cancer from medico-administrative databases?  
Eur J Epidemiol. 2008;23(10):681-8.**
- 117 Bernatsky S, Joseph L, Pineau CA, Bélisle P, Boivin JF, Banerjee D, Clarke AE**

**Estimating the prevalence of polymyositis and dermatomyositis from administrative data: age, sex, and regional differences**  
**Ann Rheum Dis. 2008 Aug 19**

- 118 San Román Terán CM, Guijarro Merino R, Guil García M, Villar Jiménez J, Martín Pérez M, Gómez Huelgas R; Efficiency Group of the Internal Medicine Services; Andalusian Society of Internal Medicine (SADEMI)  
Analysis of 27,248 hospital discharges for heart failure: a study of an administrative database 1998 – 2002  
Rev Clin Esp. 2008 Jun;208(6):281-7
- 119 Saposnik G, Cote R, Phillips S, Gubitz G, Bayer N, Minuk J, Black S; Stroke Outcome Research Canada (SORCan) Working Group  
Stroke outcome in those over 80: a multicenter cohort study across Canada  
Stroke. 2008 Aug;39(8):2310-7.
- 120 Wierdsma AI, Sytema S, van Os JJ, Mulder CL  
Case registers in psychiatry: do they still have a role for research and service monitoring?  
Curr Opin Psychiatry. 2008 Jul;21(4):379-84
- 121 Moore L, Clark DE  
The value of trauma registries  
Injury. 2008 Jun;39(6):686-95
- 122 Varas-Lorenzo C, Castellsague J, Stang MR, Tomas L, Aguado J, Perez-Gutthann S**  
**Positive predictive value of ICD-9 codes 410 and 411 in the identification of cases of acute coronary syndromes in the Saskatchewan Hospital automated database**  
**Pharmacoepidemiol Drug Saf. 2008 Aug;17(8):842-52**
- 123 Morrato EH, Dillon P, Ziegler M  
Surgical outcomes research: a progression from performance audits, to assessment of administrative databases, to prospective risk-adjusted analysis - how far have we come?  
Curr Opin Pediatr. 2008 Jun;20(3):320-5
- 124 Bahrami S, Holstein J, Chatellier G, Le Roux YE, Dormont B  
Using administrative data to assess the impact of length of stay on readmissions: study of two procedures in surgery and obstetrics  
Rev Epidemiol Sante Publique. 2008 Apr;56(2):79-85
- 125 Gerkens S, Beguin C, Crott R, Closon MC, Horsmans Y  
Assessing the quality of pharmacological treatments from administrative databases: the case of low-molecular-weight heparin after major orthopaedic surgery  
J Eval Clin Pract. 2008 Aug;14(4):585-94.
- 126 Clarke P, Leal J, Kelman C, Smith M, Colagiuri S  
Estimating the cost of complications of diabetes in Australia using administrative health-care data  
Value Health. 2008 Mar-Apr;11(2):199-206
- 127 Hser YI, Evans E  
Cross-system data linkage for treatment outcome evaluation: lessons learned from the California Treatment Outcome Project  
Eval Program Plann. 2008 May;31(2):125-35.
- 128 Prosser RJ, Carleton BC, Smith MA.**

**Identifying persons with treated asthma using administrative data via latent class modelling**

**Health Serv Res. 2008 Apr;43(2):733-54**

- 129 Escobar GJ, Fireman BH, Palen TE, Gardner MN, Lee JY, Clark MP, Kipnis P  
Risk adjusting community-acquired pneumonia hospital outcomes using automated databases  
Am J Manag Care. 2008 Mar;14(3):158-66
- 130 Bottle A, Aylin P  
Intelligent information: a national system for monitoring clinical performance  
Health Serv Res. 2008 Feb;43(1 Pt 1):10-31.
- 131 Trerotoli P, Bartolomeo N, Moretti AM, Serio G  
Hospitalisation for COPD in Puglia: the role of hospital discharge database to estimate prevalence and incidence  
Monaldi Arch Chest Dis. 2008 Sep;69(3):94-106
- 132 Haupt D, Wettermark B, Nilsson JL**  
**Dispensed volumes of anti-asthmatic drugs related to the prevalence of asthma and COPD in Sweden**  
**Pharmacoepidemiol Drug Saf. 2008 May;17(5):461-7**
- 133 Li B, Evans D, Faris P, Dean S, Quan H  
Risk adjustment performance of Charlson and Elixhauser comorbidities in ICD-9 and ICD-10 administrative databases.  
BMC Health Serv Res. 2008 Jan 14;8:12
- 134 Rasmussen S, Abildstrom SZ, Rasmussen JN, Gislason GH, Schramm TK, Folke F, Køber L, Torp-Pedersen C, Madsen M  
Hospital variation in use of secondary preventive medicine after discharge for first acute myocardial infarction during 1995-2004  
Med Care. 2008; 46: 70 – 7
- 135 Pilote L, Beck CA, Eisenberg MJ, Humphries K, Joseph L, Penrod JR, Tu JV.  
Comparing invasive and noninvasive management strategies for acute myocardial infarction using administrative databases  
Am Heart J. 2008; 155: 42 – 8
- 136 Ceratti A, Roger France FH, Beguin C  
Health data quality improvement by comparing administrative medical data and billing data.  
Int J Med Inform. 2008 Aug;77(8):527-33
- 137 Roos LL, Brownell M, Lix L, Roos NP, Walld R, MacWilliam L  
From health research to social research: privacy, methods, approaches  
Soc Sci Med. 2008; 66: 117 – 29.
- 138 Jeffrey R. Curtis, Amy S. Mudano, Daniel H. Solomon, Juan Xi, Mary Elkins Melton, and Kenneth G. Saag  
Identification and Validation of Vertebral Compression Fractures Using Administrative Claims Data  
Med Care 2009; 47: 69–72
- 139 Bruno Hubert MD, Rodica Gilca MD  
Use administrative databases with caution  
DOI:10.1503/cmaj.1080124

- 140 Spencer E. Harpe, Pharm D., Ph. D., M. P. H.  
Using secondary data sources for pharmacoepidemiology and outcomes research  
Pharmacotherapy. 2009; 29: 138 – 53
- 141 Welke KF, Diggs BS, Karamlou T, Ungerleider RM  
Comparison of pediatric cardiac surgical mortality rates from national administrative data to contemporary clinical standards  
Ann Thorac Surg. 2009; 87: 216 – 22; discussion 222-3
- 142 Szumski NR, Cheng EM**  
**Optimizing algorithms to identify Parkinson's disease cases within an administrative database**  
**Mov Disord. 2009; 24: 51 – 6**
- 143 Hanley JA, Dendukuri N  
Efficient sampling approaches to address confounding in database studies  
Stat Methods Med Res. 2009; 18: 81 – 105.
- 144 Vink NM, Klungel OH, Stolk RP, Denig P  
Comparison of various measures for assessing medication refill adherence using prescription data  
Pharmacoepidemiol Drug Saf. 2009; 18: 159 – 65
- 145 Tu K, Campbell N, Chen X-L, Cauch-Dudek K, McAlister F.**  
**Accuracy of administrative databases in identifying patients with hypertension**  
**Open Medicine. 2007; 1: 18-26.**
- 146 Soriano JB, Maier WC, Visick G, Pride NB**  
**Validation of general practitioner-diagnosed COPD in the UK general practise research database**  
**Eur J Epidemiol 2001; 17: 1075-1080**
- 147 Hansel A, Hollowell J, McNiece R, Nichols T, Strachan D**  
**Validity and interpretation of mortality, health service and survey data on COPD and asthma in England**  
**Eur Respir J 2003; 21: 279-286**
- 148 Merry AHH, Boer JMA, Schouten LJ, Feskens EJM, Verschuren WMM, Gorgels APM, van den Brandt PA**  
**Validity of coronary heart diseases and heart failure based on hospital discharge and mortality data in the Netherlands using the cardiovascular registry Maastricht cohort study**  
**Eur J Epidemiol 2009; 24:237–247**
- 149 Dendukuri N, Joseph L**  
**Bayesian approaches to modeling the conditional dependence between multiple diagnostic tests.**  
**Biometrics 2001; 57: 158 – 67**
- 150 Qu Y, Tan M, Kutner MH**  
**Random effects models in latent class analysis for evaluating accuracy of diagnostic tests**  
**Biometrics 1996; 52: 797 – 810**
- 151 Koek HL, de Bruin A, Gast A, Gevers E, Kardaun JW, Reitsma JB, Grobbee DE, Bots ML**

**Incidence of first acute myocardial infarction in the Netherlands  
Neth J Med. 2007; 65: 434 – 41**

- 152 Hughes AJ, Daniel SE, Ben Shlomo Y, Lees AJ.  
The accuracy of parkinsonian syndrome in specialist movement disorder service.  
Brain 2002; 125: 861-870.
- 153 Bernatsky s, Joseph L, Brelisle P, Boivin J-F, Rajan R, Moores A, Clarke A  
Bayesian modelling of imperfect ascertainment methods in cancer studies  
Statist Med 2005; 24: 2365 – 79**

## CONSIDERAZIONI FINALI

Il percorso che ha portato questo progetto a esplorare presso alcune ASL la possibilità di replicare un sistema di monitoraggio delle principali malattie sfruttando le enormi potenzialità che offre l'incrocio di fonti informative sanitarie elettroniche ha raggiunto risultati molto soddisfacenti.

In tutte tre le ASL coinvolte si è riusciti infatti ad attivare un Sistema Epidemiologico Integrato (SEI) in grado di fornire informazioni tempestive sulle frequenze di alcune delle maggiori patologie a livello di popolazione fornendo in tal modo la base indispensabile sia per la conoscenza epidemiologica sullo stato di salute delle popolazione che per la programmazione degli interventi sanitari di promozione della salute.

Solo da pochi anni le fonti informative sono compresenti nella gestione amministrativa delle ASL con sufficiente completezza e qualità costituendo quindi attualmente un limite all'analisi delle tendenze temporali, mentre sin d'ora i risultati sono disponibili per analisi spaziali della distribuzione delle malattie.

Lo strumento fornito e testato all'interno del progetto va considerato soprattutto in una logica di tipo prospettico nella quale l'intera popolazione degli assistibili può essere seguita nel tempo per indagini valutative, eziologiche o semplicemente di monitoraggio delle tendenze epidemiologiche costituendo così la base decisionale dei vari, complessi aspetti del governo della salute pubblica.

Questa esperienza ha anche dimostrato l'importanza della parallela e indispensabile presenza di un gruppo di esperti che sia in grado di operare criticamente in questo campo in rapida espansione e attualmente carente di basi metodologiche consolidate.

Padova, 23 Aprile 2010

Il responsabile scientifico del progetto

Lorenzo Simonato