

Le fonti dati ed i metodi di analisi

Matilde Razzanelli

Firenze, 23 maggio 2018

Sala delle fanciulle - Villa la Quiete

PROGRAMMAZIONE



Servizi

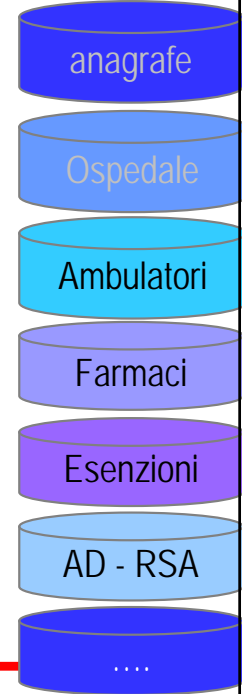
Anagrafe assistibili	Esenzioni
Ospedale	AD - RSA
Ambulatori	Riabilitazione
Farmaci	Pronto soccorso



REGIONE
TOSCANA



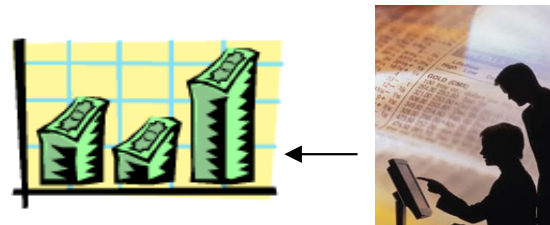
D
B
G
E
S
T
I
O
N
A
L
E



Organi
Decisori



MMG



Analisi gestionali



Debiti ministeriali

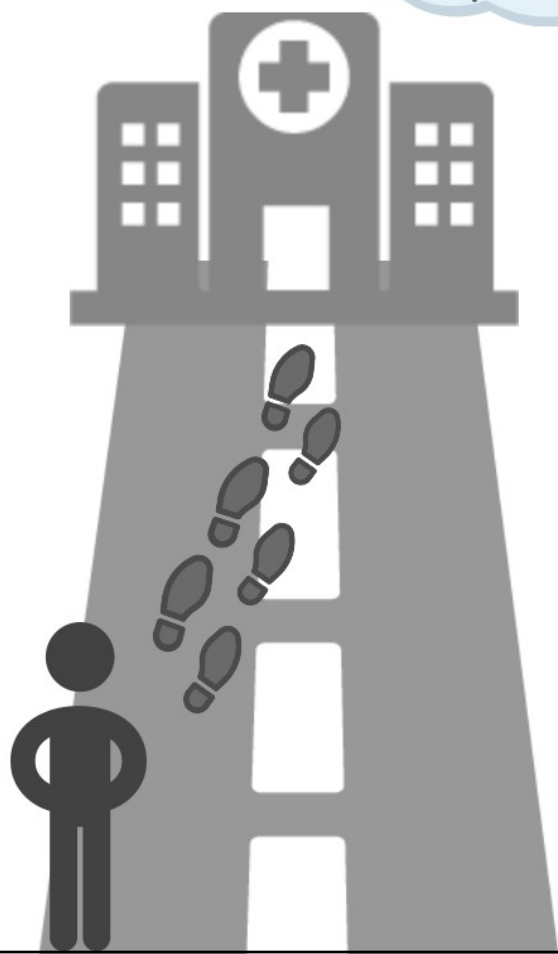
FLUSSI SANITARI TOSCANI



I cittadini toscani producono

TRACCE INFORMATICHE

quando effettuano una
qualsiasi prestazione
sanitaria pubblica o privata
convenzionata



ARS TOSCANA
agenzia regionale di sanità

FLUSSI SANITARI TOSCANI



I **flussi sanitari correnti** sono documenti elettronici (file) che descrivono in modo standardizzato una specifica categoria di eventi (ad esempio: ricoveri ospedalieri, prestazioni specialistiche, erogazioni di farmaci). Un flusso viene regolamentato da una Delibera e viene garantito un modelli di scambio tra strutture diverse con carattere di continuità e standardizzazione



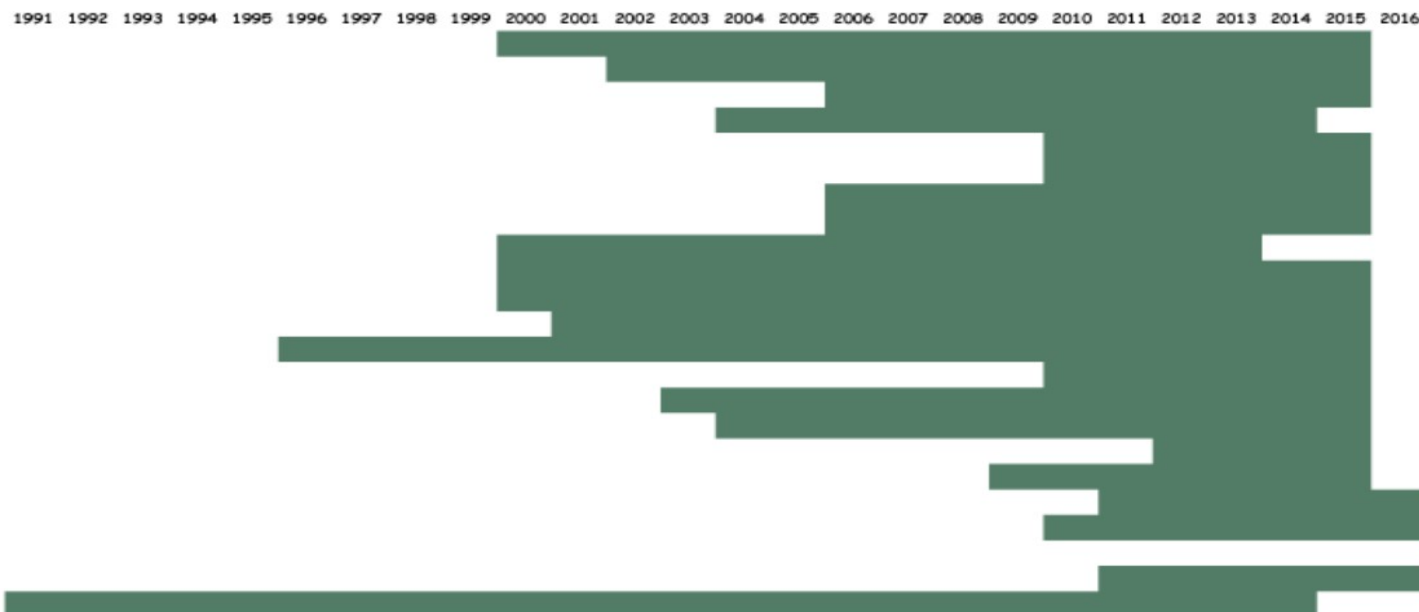
Il database ARS



Le informazioni archiviate nel DB ARS coprono un orizzonte temporale molto esteso

FLUSSO INFORMATIVO

Prestazioni riabilitazione - SPR
Specialistica ambulatoriale - SPA
Prestazioni/azioni Consultoriali - SPC
Prestazioni di assistenza protesica - SPP
Salute mentale - SALM
Archivio esenzioni - SEA
Percorsi assistenziali specialistici - PAS
Consultori - CON
Mortalità - RMR
Aborti spontanei - ABS
Interruzione volontaria di gravidanza - IVG
Certificato assistenza al parto - CAP
Scheda nosologica - SDO
Farmaci erogati dalle strutture ospedaliere - FES
Farmaceutica (ricetta SSR) - SPF
Farmaci erogati direttamente - FED
Dispositivi medici Erogati dalle Strutture - DES
Tempi di attesa - TAT
118
Accessi al pronto soccorso - PS
Assistenza domiciliare, residenziali, semiresidenziali - AD/RSA
Schede di morte informatizzate - ISM
Incidenti stradali



16 anni Prestazioni ambulatoriali, visite, diagnostica, analisi di laboratorio

22 anni dimissioni ospedaliere

15 anni di erogazioni di farmaci dalle farmacie

14 anni di farmaci erogati direttamente dalle strutture

Il lavoro sui dati

Data sources

SDO.txt
SPA.txt
SPF.txt
CAP.txt
IVG.txt
AS.txt
FED.txt
DES.txt
SPP.txt
...



ETL

caricamento file
pulizia dati
storicizzazione

2 terabyte

200 tabelle

50 utenti



Data Manipulation

calcolo indici

normalizzazione
indirizzi

campi ricodificati

criptazione iduni:
IDARS

Front-end Analytics



Data Mining

Reports

Data
Visualization

Web Portals

In informatica Extract, Transform, Load (**ETL**) è un'espressione in lingua inglese che si riferisce al processo di estrazione, trasformazione e caricamento dei dati in un sistema di sintesi (data warehouse, data mart...).



ARS TOSCANA
agenzia regionale di sanità

ARS analizza il Database storicizzato ed anonimizzato

ricostruisce i

PERCORSI DI SALUTE

dei cittadini toscani



In 1 anno 3,7 milioni di cittadini lasciano
190 milioni di tracce



ARS TOSCANA
agenzia regionale di sanità

Metodi di analisi

- Definizione e identificazione tramite dati amministrativi di coorti di osservazione
 - Popolazione generale
 - Soggetti in carico a determinati servizi
 - Coorti di patologia
 -
- Definizione e calcolo di indicatori
- Analisi e confronto degli indicatori nel tempo e nello spazio

Popolazione Assistibili

- Anagrafe assistiti Regione Toscana
 - Informazioni anagrafiche (sesso, data nascita, data morte)
 - Associazione assistito – comune di residenza
 - Associazione assistito – MMG
 - Associazione assistito - cittadinanza
- Anagrafe operatori
 - Operatori, incarico, asl di lavoro
- Tabella associazione MMG – AFT

In base a queste info ricostruiamo la popolazione assistibili, stratificandola

- Per area di **RESIDENZA** dell'assistito (MACRO LTC Riabilitazione)
o
- Per **AFT** del MMG di scelta (AFT)

Coorti di patologia



ARS TOSCANA
agenzia regionale di sanità

Regione Toscana



LA BANCA DATI DELLE MALATTIE CRONICHE AGGIORNAMENTO 2018

Documenti
ARS Toscana

marzo
2018 **99**

Col supporto dei **clinici** e della **letteratura** abbiamo definito algoritmi per l'identificazione nei dati amministrativi di coorti di pazienti affetti da una determinata patologia.

Algoritmo SEMPLICE:

→ICTUS: monoflusso: Presenza nella diagnosi primaria di un ricovero di uno dei seguenti codici: 430, 431, 432, 434, 436

Algoritmo COMPLESSO

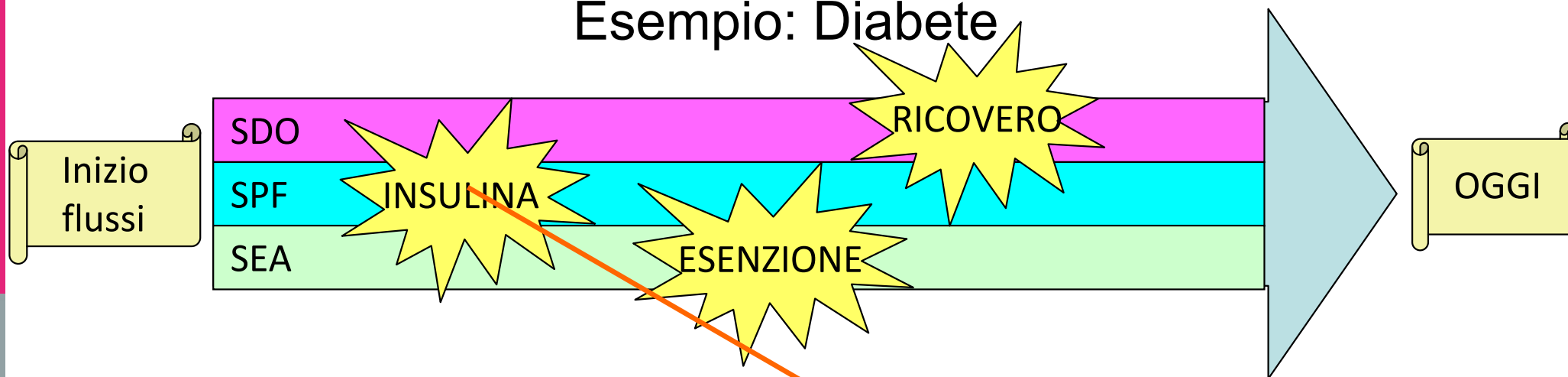
→DIADETE: interflusso: assunzione di farmaci Q ricovero Q Esenzione



ARS TOSCANA
agenzia regionale di sanità

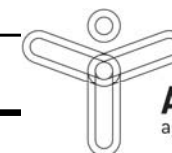
Algoritmi di prevalenza

Esempio: Diabete

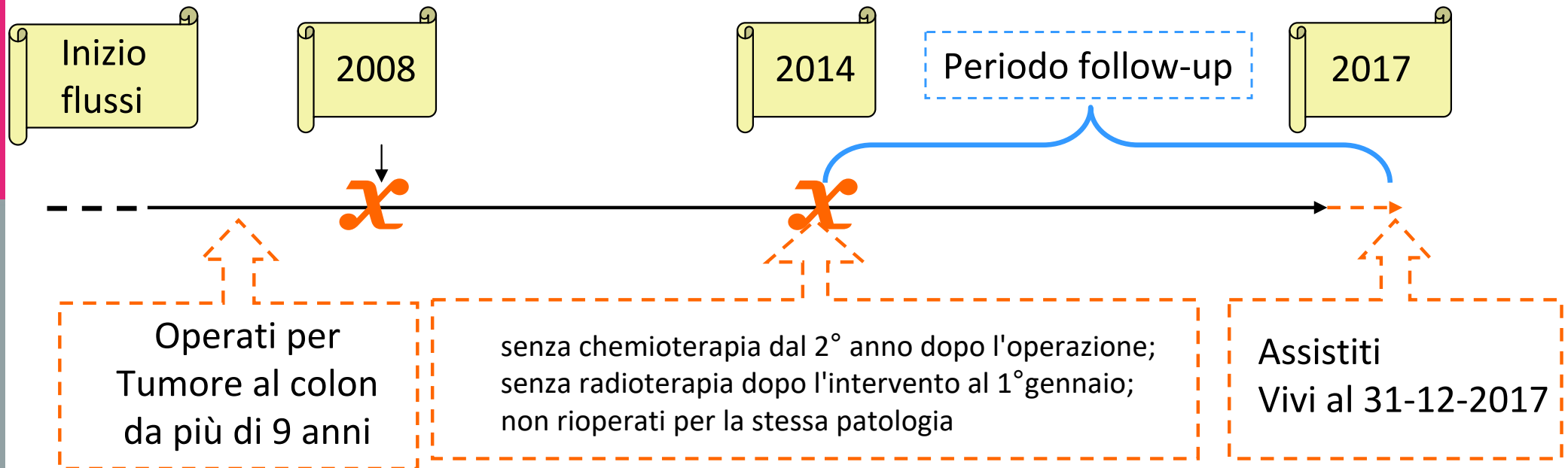


IDUNI	PATOLOGIA	DATA INIZIO
REG0000237R2Y74Y	diabete	11/05/2004
....

BPCO	Diabete
Parkinson	Cardiopatia Ischemica
Scopenso Cardiaco	Ipertensione
Ictus	Demenze
Sclerosi Multipla	IMA



Algoritmi di follow-up: coorte Follow-up Tumori 2017



In sintesi

- Si catturano dalle dimissioni ospedaliere (SDO) tutti gli interventi per tumore al colon
- Si filtrano tra questi i soggetti vivi alla fine del periodo di osservazione
- Si escludono tutti quelli che hanno fatto chemioterapia o radioterapia o hanno subito un re-intervento tra l'intervento indice e la data di osservazione
- Si controlla nella coorte quanti hanno fatto una colonscopia tra il 6 e il 10 anno dall'intervento

Indicatori

- Gli indicatori (consumo, processo, esito ecc.) sono stati catalogati e raggruppati per fonte (flusso/flussi di riferimento)
- Ogni gruppo ha una sua procedura di estrazione primaria di dati che raccoglie un dataset individuale
- Che costituisce la base per il successivo calcolo dei vari livelli di aggregazione
- In seguito vengono indirizzati a procedure di pubblicazione (WEB, PDF ecc.)

In pratica...traduciamo un algoritmo

7. MORBO DI PARKINSON

7.1 ALGORITMO DI IDENTIFICAZIONE

I pazienti vengono estratti se sussiste almeno una tra le seguenti condizioni in tabella 7.1.

Tabella 7.1
Criteri per l'identificazione del morbo di Parkinson

CODICI	DESCRIZIONE
SDO	Presenza in una qualsiasi delle diagnosi di dimissione principale o secondaria con codice 332*, (ovvero 332 - Morbo di Parkinson, 3320 - Paralisi agitante, 3321 - Parkinsonismo secondario)
SPF o FED	Presenza di almeno due prescrizioni in date distinte codice ATC5 = N04BA03 (Levodopa inibitore della decarbossilasi e della Comt), N04BC07 (Apomorfina), N04BC09 (Rotigotina), N04BD02 (Rasagilina), N04BX01 (Tolcapone), N04BX02 (Entacapone); oppure con codice prestazione = 0032261149, 0032261152, 0032261164, 0032261176, 0032261240, 0032261265, 0034090035, 0034090264, 0034090050, 0034090062, 0034090074, 0034090086, 0034090098, 0034090100, 0034090163, 0034090201, 0034090237, 0025462019, 0025462021. Questi codici prestazione, a differenza degli ATC5, identificano più precisamente la molecola e il dosaggio e corrispondono ai farmaci elencati nella tabella A.6 nel materiale supplementare
SEA	Presenza di un'esenzione con codice 332
AD-RSA	Patologia prevalente o concomitante con codice ICD9CM 332 o presenza di Parkinson o Parkinson come causa principale di disabilità (7)

Nel procedimento necessario ad estrarlo dai dati disponibili

```
• create or replace
• PROCEDURE          MACREX_00_ADM(data_indice varchar2) IS

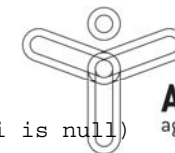
• --PROCEDURA CHE GESTISCE L'ESTRAZIONE DELLA BACA DATI MACRO
• ANNO varchar2(20);
• OGGI VARCHAR2(20);
• [...]
• /*aggiorna tabella annuale fibrillazione atriale in base alla data indice*/
• MACREX_05_FIB_ATRIALE(ANNO);

• NOME_PARTIZIONE := REPLACE(DATA_INDICE,'/', '');
• SELECT COUNT(*) INTO NRO_RECORD FROM USER_TAB_PARTITIONS where table_name = 'MACREX_DATASET_ANNO' AND
PARTITION_NAME='PARTIZIONE_'||NOME_PARTIZIONE||'';
• IF (NRO_RECORD>0) THEN
• EXECUTE IMMEDIATE 'alter table MACREX_DATASET_ANNO drop partition PARTIZIONE_'||NOME_PARTIZIONE||'';
• END IF;
• EXECUTE IMMEDIATE 'alter table MACREX_DATASET_ANNO add partition PARTIZIONE_'||NOME_PARTIZIONE||' VALUES
(''||DATA_INDICE||''') STORAGE (INITIAL 140M NEXT 14M) TABLESPACE FLUSSI '; COMMIT;

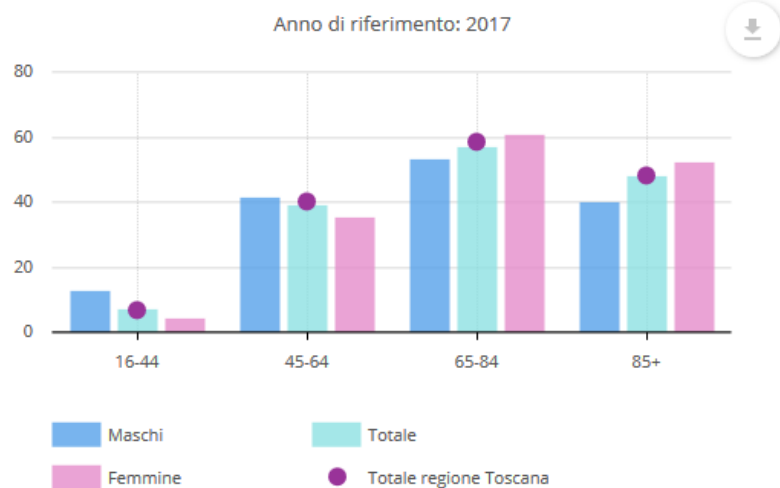
• DECLARE
• CURSOR CURSORE_macro IS
• SELECT patologia from macrex_dataset_prevalenti group by patologia;

• BEGIN
• OPEN CURSORE_MACRO;
• LOOP
• FETCH CURSORE_MACRO INTO v_patologia;
• EXIT WHEN CURSORE_MACRO%NOTFOUND;

• DELETE FROM MACREX_DATASET_ANNO WHERE INDICATORE=v_patologia AND DATA_INDICE=P_DATA_INDICE; COMMIT;
• COMANDO:='
• INSERT INTO MACREX_DATASET_ANNO ( IDUNI, DATA_INDICE, INDICATORE, NUMERATORE, DATA_AGG)
• select ID_UNIVERSALE , ''||P_DATA_INDICE||'' DATA_INDICE, 'MACRO_'||''||v_patologia||'', '1' NUMERATORE ,
• ''||OGGI||'' DATA_AGG
• FROM
• (SELECT IDUNI
• FROM derivars.MACREX_DATASET_PREVALENTI MACRO
• WHERE patologia='''||v_patologia||''
• AND DATA_INIZIO<= TO_DATE ('''||P_DATA_INDICE||'', 'DD/MM/YYYY'))
• ) MACRO,
• ANAGRAFE.U_ARS_anag_med_res_storico ANA1
• WHERE ANA1.ini_record<=TO_DATE('''||P_DATA_INDICE||'', 'DD/MM/YYYY')
• AND ANA1.fine_record>TO_DATE('''||P_DATA_INDICE||'', 'DD/MM/YYYY')
• and (ana1.data_morte_marsi>TO_DATE('''||P_DATA_INDICE||'', 'DD/MM/YYYY') or ana1.data_morte_marsi is null)
• AND MACRO.IDUNI=ANA1.ID_UNIVERSALE
```

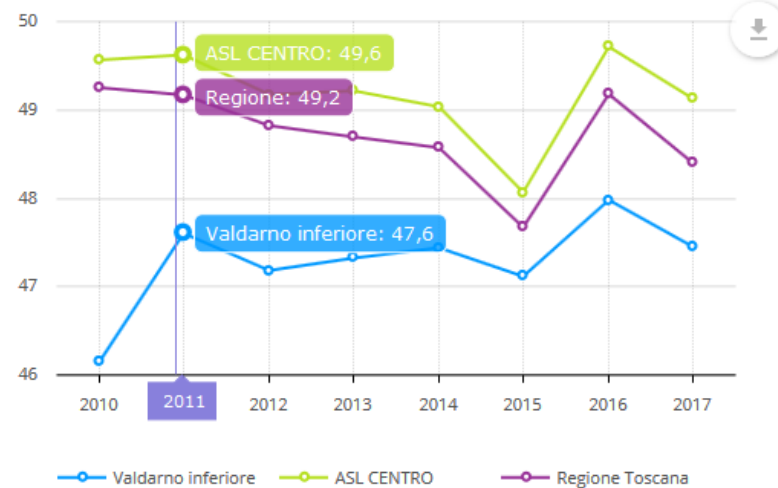


Prevalenza per sesso e fascia d'età



[Guida alla lettura](#)

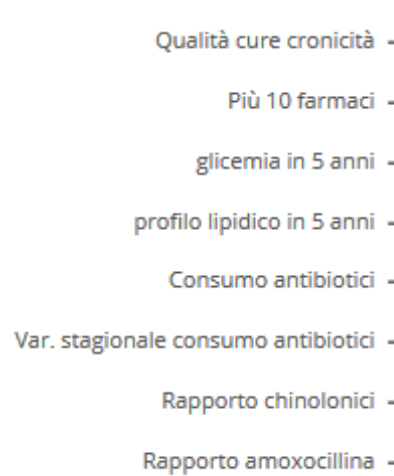
Andamento temporale



[Guida alla lettura](#)

Assistiti ultra15enni

AFT	REGIONE
32	32
2	2,3
42	42,5
35,9	37,8
17	18,2
40,8	41,2
14,5	14,2
15,4	12,7



- Punti di forza

- Enorme mole di dati disponibili
- Disponibilità di un lungo periodo di osservazione
- Procedure standardizzate di estrazione e calcolo

- Criticità

- I dati amministrativi rappresentano la maggior parte delle cure erogate, ma non TUTTE (vedi specialistica ambulatoriale e esami di laboratorio¹)
- Qualità dei dati non sempre ottimale
- Carenza di dati clinici

Grazie per l'attenzione