

Alla ricerca delle fonti



Roberto Raschetti

Roma, Istituto Superiore di Sanità

15 giugno 2017

Accedere ai dati sull'uso dei farmaci

Quesito di ricerca

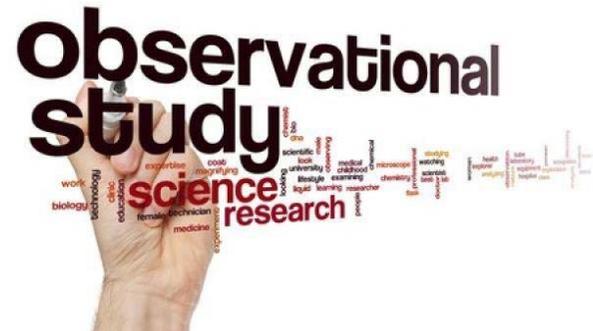


Quali dati sono necessari?



Come, dove acquisire i dati?

tempi
costi
fattibilità



Tipi di dati

Dati primari

Dati secondari





ELSEVIER

Journal of Clinical Epidemiology 66 (2013) 703–705

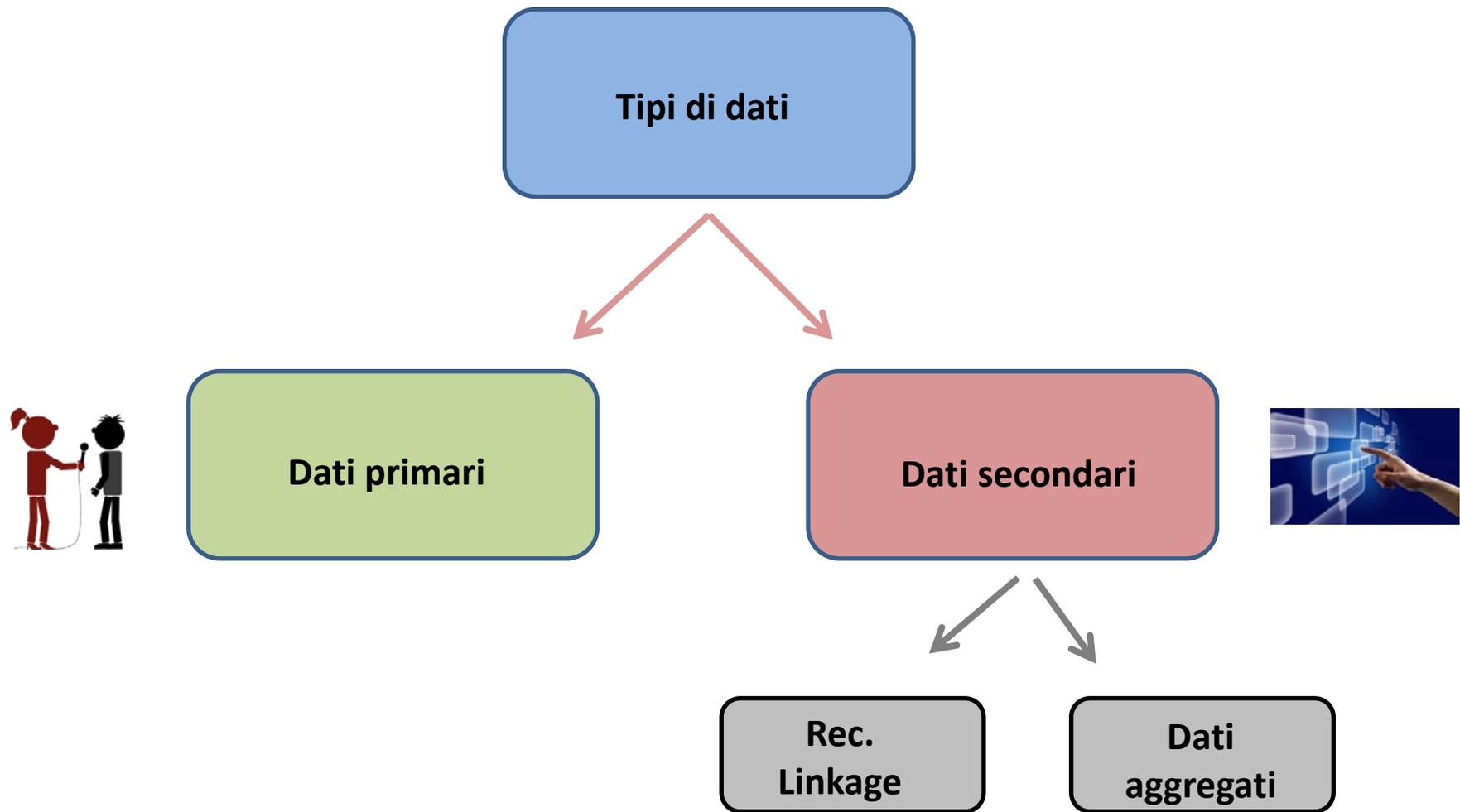
**Journal of
Clinical
Epidemiology**

COMMENTARIES

**Call to RECORD: the need for complete reporting of research using
routinely collected health data**

Eric I. Benchimol^{a,b,c,*}, Sinead Langan^d, Astrid Guttman^{a,e,f}, on behalf of the
RECORD Steering Committee¹

The rapid entry of health care into the electronic age has led to a proliferation of large data repositories of health-related information. Such data are collected for administrative and clinical purposes, without specific *a priori* research questions. Examples of these “routinely collected health data” include health administrative data, data warehouses of electronic medical records, primary care medical record data, and disease registries. These data are increasingly being used for observational, comparative effectiveness and health services research.



WORKSHOP

“Quale sistema informativo sanitario per la sostenibilità e lo sviluppo del Servizio Sanitario Nazionale”

Istituto Superiore di Sanità
8 giugno 2016



Sessione 3 : Accessibilità dei dati raccolti nel SSN e tutela della privacy

I **dati aggregati** (le combinazioni di modalità alle quali è associata una frequenza non inferiore a una soglia prestabilita) non presentano vincoli imposti a protezione del dato personale. Ma, non sempre i dati aggregati possono realizzare le conoscenze che la ricerca epidemiologica e biomedica ricercano.

Rec Linkage: messa a punto di sistemi informatici di “*anonimizzazione reversibile*”, cioè di sistemi che dopo la fase di controllo del dato permettano l’anonimizzazione del dato con un codice.

Tuttavia al momento non è disponibile un sistema di anonimizzazione nazionale dei dati sanitari e i sistemi di anonimizzazione a livello regionale hanno coperture difformi per tipo di flussi interessati e arco temporale coperto.



Accessibilità



Trasparenza

Accountability





Codice dell'amministrazione digitale

aggiornato al D.Lgs. 179/2016

Lo Stato, le Regioni e le autonomie locali assicurano **la disponibilità, la gestione, l'accesso, la trasmissione, la conservazione e la fruibilità** dell'informazione in modalità digitale e si organizzano ed agiscono a tale fine utilizzando con le modalità più appropriate le tecnologie dell'informazione e della comunicazione.

(art. 2, comma 1, D. lgs. n. 82/2005)

I dati delle Pubbliche Amministrazioni sono **formati, raccolti, conservati, resi disponibili e accessibili** con l'uso delle tecnologie dell'informazione e della comunicazione che ne consentano la fruizione e riutilizzo, **alle condizioni fissate dall'ordinamento**, da parte delle altre Pubbliche Amministrazioni e dai privati.

(art. 50, comma 1, D. lgs. n. 82/2005)

Open data

I dati aperti sono dati che possono essere liberamente utilizzati, riutilizzati e ridistribuiti da chiunque, eventualmente soggetti alla necessità di citarne la fonte e di condividerli con lo stesso tipo di licenza con cui sono stati originariamente rilasciati.



Livelli di opennness

Una scala a cinque stelle

OL : Open Licence



Il dato è disponibile sul web (in qualsiasi formato) con una licenza aperta

RE : REadable



dati strutturati ma codificati con un formato proprietario: ad esempio Microsoft Excel

OF : Open Format



dati strutturati e codificati in un formato non proprietario: ad esempio il formato .csv (*Comma Separated Values*)

UR : Uniform Resource identifier



[http://data....](http://data...)

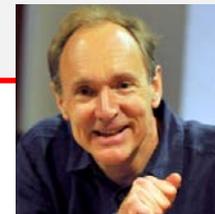
dati strutturati e codificati in un formato non proprietario che sono dotati di un URI (*Identificatore Univoco di Risorsa*) che li rende indirizzabili sulla rete e quindi utilizzabili direttamente online

LD : Linked Data



Il dato rispetta tutti gli altri criteri e inoltre contiene collegamenti ad altri dati (W3C: Web semantico)

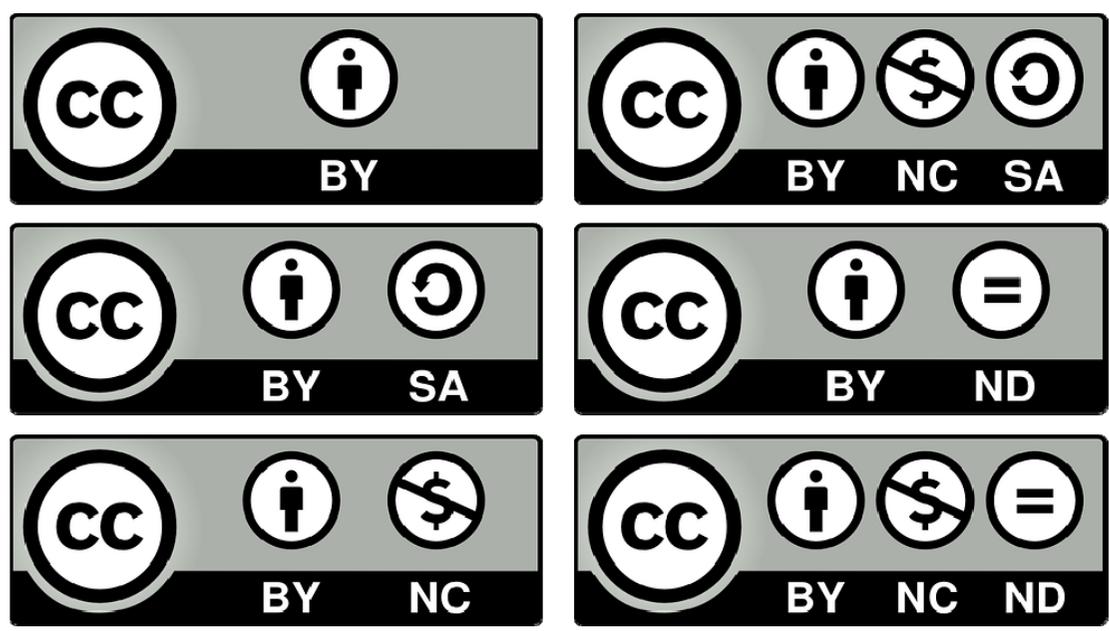
Tim Berners Lee (inventore del World Wide Web)



Modalità di utilizzo

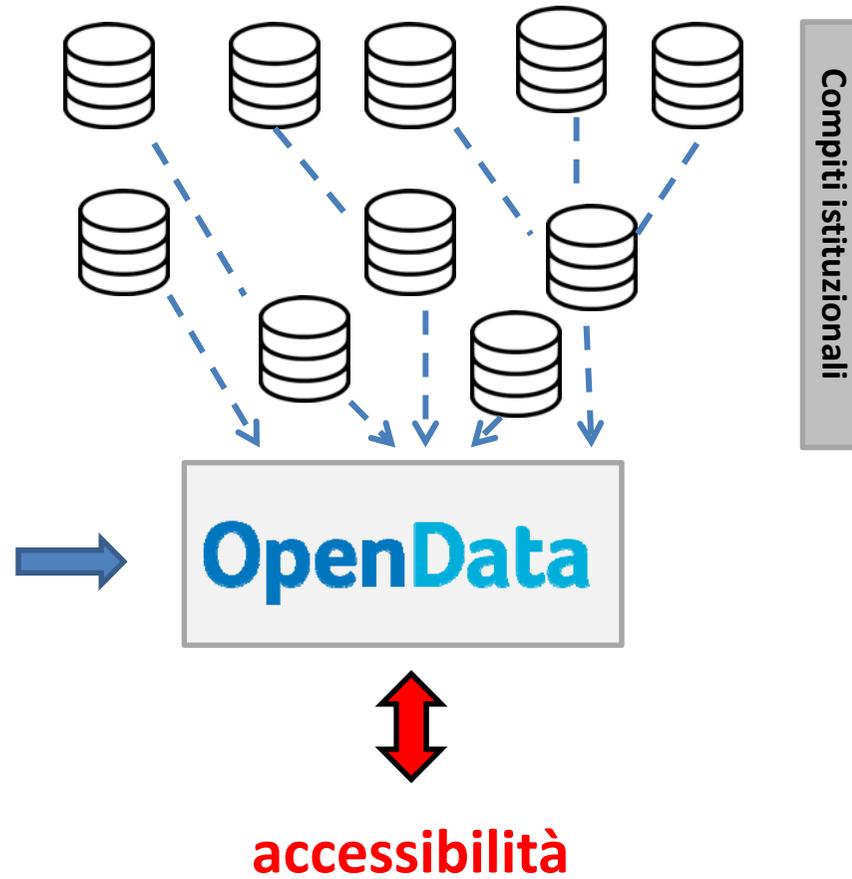


Licenze Creative Commons



- Attribuzione (BY)
- Non commerciale (NC)
- Non opere derivate (ND)
- Condividi allo stesso modo (SA)

datawarehouse



Vincoli:

- ♣ segreto di stato
- ♣ segreto statistico
- ♣ diritto d'autore
- ♣ privacy
- ♣ sicurezza pubblica
- ♣ ecc

Alcuni esempi

FDA U.S. FOOD & DRUG ADMINISTRATION A to Z Index | Follow FDA | En Español

Search FDA

Home | Food | Drugs | Medical Devices | Radiation-Emitting Products | Vaccines, Blood & Biologics | Animal & Veterinary | Cosmetics | Tobacco Products

Drugs

Home > Drugs > Guidance, Compliance & Regulatory Information > Surveillance > FDA Adverse Events Reporting System (FAERS)

FDA Adverse Events Reporting System (FAERS)

- ▶ [FDA Adverse Event Reporting System \(FAERS\): Latest Quarterly Data Files](#)
- [FDA Adverse Events Reporting System \(FAERS\) Public Dashboard](#)
- [Potential Signals of Serious Risks/New Safety Information Identified from the FDA Adverse Event Reporting System \(FAERS\)](#)
- [FDA Adverse Events Reporting System \(FAERS\) Electronic Submissions](#)

FDA Adverse Event Reporting System (FAERS): Latest Quarterly Data Files

[SHARE](#)
[TWEET](#)
[LINKEDIN](#)
[PIN IT](#)
[EMAIL](#)
[PRINT](#)

The files listed on this page contain raw data extracted from the AERS database for the indicated time ranges and are not cumulative.

Users of these files need to be familiar with creation of relational databases using applications such as ORACLE®, Microsoft Office Access, MySQL® and IBM DB2 or the use of ASCII files with SAS® analytic tools.

A simple search of FAERS data cannot be performed with these files by persons who are not familiar with creation of relational databases. However, you can get a summary FAERS report for a product by [sending a Freedom of Information Act \(FOIA\) request](#) to FDA. You can also request individual case reports by [submitting a FOIA request](#) listing case report numbers. You can get older AERS data files from the [Quarterly Legacy AERS Data Files](#) page.

The quarterly data files, which are available in ASCII or SGML formats, include:

- demographic and administrative information and the initial report image ID number (if available);
- drug information from the case reports;
- reaction information from the reports;
- patient outcome information from the reports;
- information on the source of the reports;
- a "README" file containing a description of the files.



Public Domain

FAERS Data Files

Click on a Link Below to Begin Downloading

- [FAERS ASCII 2016q4 \(ZIP - 39.2MB\)](#)
October - December 2016
- [FAERS XML 2016q4 \(ZIP - 72.3MB\)](#)
October - December 2016

primaryid\$caseid\$caseversion\$si_f_code\$event_dt\$mfr_dt\$init_fda_dt\$fda_dt\$rept_cod\$auth_num\$mfr_num\$mfr_sndr\$lit_ref\$age\$age_cod\$age_g
2\$10005259\$2\$F\$20130809\$20161202\$20140312\$20161209\$EXP\$PHHY2014BR029413\$NOVARTIS\$\$35.07187\$YR\$\$F\$Y\$60\$KG\$20161209\$\$MD\$BR\$BR100
100101343\$10010134\$3\$F\$\$20161102\$20140312\$20161114\$PER\$\$US-INCYTE CORPORATION-2013IN001384\$INCYTE\$\$\$\$F\$Y\$\$20161114\$\$MD\$US\$US1
8\$22\$F\$20131119\$20161228\$20140315\$20161230\$EXP\$CA-ROCHE-1365636\$ROCHE\$\$53.70568\$YR\$\$M\$Y\$72\$KG\$20161230\$\$OT\$CA\$CA100139636\$100:
;8 (9):1138-9.\$60\$YR\$A\$F\$Y\$\$20161028\$\$OT\$FR\$FR100151054\$10015105\$4\$F\$\$20161110\$20140317\$20161123\$EXP\$JP-PFIZER INC-2014072701\$PFIZEI
18.82\$KG\$20161007\$CN\$US\$US100181513\$10018151\$3\$F\$20110708\$20161104\$20140318\$20161110\$EXP\$JP-BEH-2011029169\$CSL BEHRING\$\$13\$YR\$T
VARTIS\$\$67.30459\$YR\$\$M\$Y\$\$20161123\$CN\$COUNTRY NOT SPECIFIED\$CA100220804\$10022080\$4\$F\$201402\$20161013\$20140319\$20161020\$EXP\$C/
0161017\$OT\$FR\$FR100273332\$10027333\$2\$F\$201401\$20160928\$20140321\$20161004\$EXP\$US-OTSUKA-US-2014-10505\$OTSUKA\$\$23.30459\$YR\$\$M\$Y;
8\$\$OT\$CA\$CA1002864810\$10028648\$10\$F\$20120521\$20161130\$20140321\$20161202\$EXP\$GB-ROCHE-1363429\$ROCHE\$\$55.45517\$YR\$\$F\$Y\$70.8\$KG\$20
\$20161014\$PER\$\$US-PFIZER INC-2014083554\$PFIZER\$\$70\$YR\$\$F\$Y\$\$20161014\$\$MD\$US\$US100318906\$10031890\$6\$F\$2012\$20161202\$20140324\$20161
1\$20\$F\$2011\$20161124\$20140326\$20161206\$EXP\$PHHY2014BR016790\$SANDOZ\$\$\$\$F\$Y\$67\$KG\$20161206\$\$MD\$BR\$BR100392327\$10039232\$7\$F\$201
213\$20140328\$20161223\$PER\$\$US-PFIZER INC-2014088237\$PFIZER\$\$46\$YR\$\$F\$Y\$122.45\$KG\$20161223\$\$MD\$US\$US100448966\$10044896\$6\$F\$2016111
IBB COMPANY-20585071\$BRISTOL MYERS SQUIBB\$\$63\$YR\$\$F\$Y\$\$20161201\$CN\$DE\$DE100477934\$10047793\$4\$F\$20131020\$20161129\$20140331\$2016
0140401\$20161204\$EXP\$US-GLAXOSMITHKLINE-A1008405A\$GLAXOSMITHKLINE\$\$\$\$F\$Y\$\$20161204\$CN\$US\$US1005157410\$10051574\$10\$F\$2003\$2
EXP\$PHHY2014CA029735\$NOVARTIS\$\$58.58\$YR\$\$M\$Y\$68.48\$KG\$20161013\$\$MD\$CA\$CA100545582\$10054558\$2\$F\$201201\$20161212\$20140402\$20161
140403\$20161031\$EXP\$PHHY2014JP037801\$SANDOZ\$\$\$\$A\$F\$Y\$\$20161031\$CN\$JP\$JP100567912\$10056791\$2\$F\$2010\$20161020\$20140403\$20161031
Y\$\$20161031\$\$OT\$JP\$JP100568002\$10056800\$2\$F\$2010\$20161020\$20140403\$20161031\$EXP\$PHHY2014JP037799\$SANDOZ\$\$\$\$A\$F\$Y\$\$20161031\$C
\$5\$F\$20131104\$20161005\$20140405\$20161010\$EXP\$CA-ROCHE-1292211\$ROCHE\$\$49.71389\$YR\$\$F\$Y\$\$20161010\$\$OT\$CA\$CA100620056\$10062005\$6;
889\$2\$F\$2011\$20160929\$20140408\$20161004\$EXP\$US-009507513-140411\$SANDOZ\$MERC\$K\$C\$C\$F\$Y\$81 63\$KG\$20161004\$MD\$US\$US100649322\$10064

The ASCII data files are '\$' delimited

Medstat.dk

On medstat.dk, you can find statistics on the total sales of medicines in Denmark 1996-2016.

The sales information is based on data from the Register of Medicinal Product Statistics.

Data for 2016 is preliminary. This is because over the counter sales outside pharmacies are not yet fully reported, hence figures for this sale are not presented. More info available [here](#).

Search modules:

- **Groups of medicines**

Groups of medicines are groups of medicines/treatments. The groups are combinations of ATC codes, product names or just a single ATC code. The shown groups only cover a smaller part of the medicine sales in Denmark and only sales from the primary sector.

- **ATC code**

The medicine sales shown are calculated by the [ATC code](#). In this search module you find the total sales of medicines in Denmark. The figures show the total sales and the sales by gender, age, region and sector.

- **Product name**

The search module "Product name" also shows the total sales of medicines however, here the sales is shown on a package level. The sales can also be divided by sector and for the primary sector we furthermore show statistics depending on whether the package is allowed for OTC sales or not.

Obs. If you are searching data for the hospital sector and ATC code L01 or J01 you are advised to read the important information in "Data basis and description" in the section "Reporting errors, deficiencies etc."

It is permitted to use the data with reference to: The Danish Health Data Authority, medstat.dk and date.

Please send your suggestions for improvement of medstat.dk to medicindata@sundhedsdata.dk

LINKS AND DOCUMENTS

[Data basis and description](#)

[ATC-system](#)

[Download data](#)

Last modified 16.03.2017

Download of metadata from medstat.dk

Here you can download metadata in search modules "ATC code" and "Product name". The data files are semi colon separated csv files. The data files and variables are described in the [download description](#). In addition to medstat data you find a data file containing population data by January 1st each year. Data is derived from [Statistics Denmark](#).

File name

[1996_atc_code_data.txt](#)
[1996_product_name_data.txt](#)
[1997_atc_code_data.txt](#)
[1997_product_name_data.txt](#)
[1998_atc_code_data.txt](#)
[1998_product_name_data.txt](#)
[1999_atc_code_data.txt](#)
[1999_product_name_data.txt](#)
[2000_atc_code_data.txt](#)
[2000_product_name_data.txt](#)
[2001_atc_code_data.txt](#)

LINKS AND DOCUMENTS

[Data basis and description](#)

[ATC-system](#)

[Download data](#)

Information

For medstat.dk we only use the latest updated DDD values and atc codes. If you have downloaded an earlier version of data you may need to download data once more to have the most updated version. Furthermore data may change between updates for other changes in The Register of Medicinal Product Statistics.

[Legal aspects of download.](#)

A01AA01;2016;100;066411;156.8;2.2;156.8;11897.7;144.1;A01AA01;2016;101;066411;156.8;156.8;11897.7;A01AA01;2016;200;066411;0.6;0.6;30.3;73955;0;0;A01AD02;2016;103;073955;0.3;1.7;16.4;A01AD02;2016;105;073955;23.4;156;1536.1;A01AD02;2016;300;073955;A01AD02;2016;01AD11;2016;101;056036;3.9;38.7;513.6;A01AD11;2016;200;056036;0;0;0.2;A01AD11;2016;300;056036;3.9;38.8;A02AA04;2016;100;004659;30;A02AA04;2016;100;019541;0;0;0;A02AA04;2016;103;019541;0;0.2;0.5;A02AA04;2016;300;019541;A02AA04;2016;100;084389;0.8;3.4;33.7;253.5;A02AD01;2016;105;040318;111.5;1115.2;8203;A02AD01;2016;200;040318;0.7;6.9;36.5;A02AD01;2016;300;040318;A02058.6;A02AD01;2016;200;396410;0.5;9.1;37.5;A02AD01;2016;300;396410;A02AD01;2016;100;407758;0;1.5;A02AD01;2016;103;4077580360;0.2;4.7;57.1;A02BA02;2016;300;040360;A02BA02;2016;100;061687;0.6;48.2;A02BA02;2016;103;061687;4.8;145;598.2;A02BA02;2016;103;495259;0.3;3.9;46.5;A02BA02;2016;105;495259;0.4;6.3;71.8;A02BA02;2016;200;495259;0;0.3;2.3;A02BA02;2016;300;495248.6;A02BB01;2016;300;003455;2.1;31.4;A02BC01;2016;100;021066;0;0;0.1;0;A02BC01;2016;101;021066;0;0;0.1;A02BC01;2016;300;0201296;10.8;10.7;602.2;341.3;113;A02BC01;2016;101;101296;10.8;602.2;341.3;A02BC01;2016;200;101296;0;0.3;0.1;A02BC01;2016;300;101296;016;100;396841;19.6;19.5;1094.9;570;210.1;A02BC01;2016;101;396841;19.6;1094.9;570;A02BC01;2016;300;396841;19.6;1094.9;A02BC01;2016;973.8;A02BC01;2016;101;488627;2.5;141.6;2319.9;A02BC01;2016;300;488627;2.5;141.6;A02BC01;2016;100;493226;28.8;0.2;403.9;982.9;2.1;A200;540432;0;0.8;0.2;A02BC01;2016;300;540432;10.6;295.8;A02BC01;2016;100;555860;2.7;2.6;160.2;564.8;200.3;A02BC01;2016;101;555860;24.2;378.3;A02BC01;2016;101;589198;25.2;2519.3;954.2;A02BC01;2016;200;589198;0;0.3;0.1;A02BC01;2016;300;589198;25.2;2519.6;A02BC00;A02BC02;2016;300;033612;15.7;440.8;A02BC02;2016;100;033623;17;17;952.3;647.9;253.4;A02BC02;2016;101;033623;17;952.3;647.9;A02BC02;1407;599.7;A02BC02;2016;101;050893;38.9;1944.2;1407;A02BC02;2016;200;050893;0.3;15.4;5;A02BC02;2016;300;050893;39.2;1959.6;A0088145;11.6;648.8;410.4;A02BC02;2016;200;088145;11.6;648.8;A02BC02;2016;100;101452;0;0.4;2.5;0.5;A02BC02;2016;101;101452;0;0.4;2.5;

Home

Attività

- > Registrazione
- > Sicurezza
- > Farmaci falsificati, illegali e rubati
- > Ispezioni
- > Negoziazione e rimborsabilità
- > Consumi e spesa farmaceutica
- > Informazione indipendente
- > Sperimentazione e ricerca
- > Registri Farmaci sottoposti a monitoraggio
- > Rapporti Internazionali
- > Affari amministrativi
- > Qualità delle procedure e Controllo di gestione
- > Amministrazione Trasparente

Open Data

Con Dati aperti si fa riferimento alla dottrina per la quale alcune tipologie di dati in possesso delle Pubbliche Amministrazioni sono rese liberamente accessibili a tutti, senza restrizioni di copyright che ne limitino la riproduzione.

Il Codice dell'Amministrazione Digitale (art. 68 c. 3 lett. b) definisce "dati aperti" quei dati che presentano le seguenti caratteristiche:

- sono disponibili secondo i termini di una licenza che ne permetta l'utilizzo da parte di chiunque, anche per finalità commerciali, in formato disaggregato;
- sono accessibili attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, in formati aperti ai sensi della lettera a), sono adatti all'utilizzo automatico da parte di programmi per elaboratori e sono provvisti dei relativi metadati;
- sono resi disponibili gratuitamente attraverso le tecnologie dell'informazione e della comunicazione, ivi comprese le reti telematiche pubbliche e private, oppure sono resi disponibili ai costi marginali sostenuti per la loro riproduzione e divulgazione.

I dati sono organizzati in un catalogo che ne permette la consultazione e il download: è sufficiente selezionare la categoria di interesse e cliccare sull'icona per effettuare lo scaricamento del dato, all'interno della scheda di dettaglio.

Non è necessaria alcuna procedura di autenticazione per effettuare il download, si consiglia tuttavia di prendere visione della licenza d'uso a corredo, in quanto questa ne determina l'uso che si può fare del dato.



La licenza di distribuzione dei dati utilizzata da AIFA la CC-BY (attribuzione) nella versione 4.0. Questa licenza permette a terzi di distribuire, modificare, ottimizzare ed utilizzare i dati, anche commercialmente, con l'obbligo di citare la fonte. Si rimanda, per maggiori informazioni, alla licenza e alla pagina di Creative Commons . I dati personali pubblicati sono riutilizzabili solo alle condizioni previste dalla normativa vigente sul riutilizzo dei dati pubblici (direttiva comunitaria 2003/98/CE e d. lgs. 36/2006 di recepimento della stessa), in termini compatibili con gli scopi per i quali sono stati raccolti e registrati, e nel rispetto della normativa in materia di protezione dei dati personali. Ulteriori informazioni nella pagina Privacy.

AIFA rende disponibili i dati per facilitarne la consultazione, il riutilizzo e la distribuzione in diversi formati (xml, csv, pdf).

Il file csv (comma separated value) è un formato testuale che consente di distribuire dati in formato tabellare: possono essere letti con software open (calc di OpenOffice) o con software proprietari come Excel di Microsoft o semplicemente con editor testuale (NotePad).

Per ulteriori informazioni sugli Open Data, consultare il Vademecum Open Data e le "Linee guida per la valorizzazione del patrimonio informativo pubblico - anno 2014"

Open data: una risorsa collettiva per combattere la corruzione e favorire la conoscenza (Sondaggio Open Data AIFA)

- [Dati su Organizzazione e Personale](#)
- [Dati su Provvedimenti AIFA](#)
- [Dati su incarichi e consulenze](#)
- [Dati su Sovvenzioni, contributi, sussidi, vantaggi economici](#)
- [Dati su Bandi di Gara e Contratti](#)
- [Dati sulle Liste dei Farmaci](#)
- [Dati sulla Farmacovigilanza](#)
- [Dati sulle Officine autorizzate](#)



[Home](#) / [Istituto nazionale di statistica](#) / [Attività](#) / [Open data in Istat](#)

[[English](#)]

Open data in Istat ASCOLTA

OPEN DATA

Con il termine **open data** si fa riferimento alla pratica di mettere a disposizione dati che possono essere liberamente utilizzati, riutilizzati e redistribuiti, con la sola limitazione, al massimo, del requisito di attribuzione e condivisione nello stesso modo (<http://opendefinition.org>).

Licenza

Tutti i dati prodotti dall'Istituto nazionale di statistica sono rilasciati sotto licenza **Creative Commons (CC BY 3.0 IT)**: è possibile riprodurre, distribuire, trasmettere e adattare liberamente dati e analisi dell'Istituto nazionale di statistica, anche a scopi commerciali, a condizione che venga citata la fonte.

Quali dati?

L'Istat è l'amministrazione che contribuisce di più in termini di dataset rilasciati: è in sostanza il maggior produttore di open data in Italia. A rilevarlo è anche dati.gov.it, il portale che si occupa dello stato dell'open data in Italia.

Sono scaricabili e rielaborabili tutti i dati dell'Istat diffusi attraverso le **banche dati** e i **sistemi informativi**; sono rilasciate in formato .xls (rielaborabile anche con software open) le tavole allegate ai **volumi** e le serie storiche allegate ai **comunicati stampa**.

Inoltre l'Istituto rende disponibili i dati del corporate datawarehouse **I.Stat** anche attraverso **web service**, grazie ad uno scambio machine-to-machine in tempo reale. Enti, organizzazioni e altri utenti possono consultare i set di dati di I.Stat e/o inserirli all'interno dei propri sistemi informativi, basi dati, portali web, ecc.

Nel maggio 2015 l'Istat ha reso disponibile La piattaforma **Linked Open Data (LOD)**, che consente di accedere e navigare dati in formato open, sulla base di tecnologie e standard del web semantico. I LOD, interrogabili direttamente da qualsiasi applicazione, rispondono alle esigenze espresse dalle comunità di utilizzatori di disporre di dati standardizzati e interoperabili.

Organizzazione

- Organigramma
- Sedi territoriali
- Normativa

Attività

- Rilevazioni ed elaborazioni
- Attività internazionali
- Promozione della cultura statistica
- [Open data in Istat](#)

Qualità

- Il percorso fatto
- Codici della statistica
- Strumenti

Presidente

- Articoli e interviste
- Interventi e relazioni

Eventi e incontri scientifici

- Conferenze di statistica
- Presentazioni del Rapporto annuale
- Giornate della statistica

Audizioni

Who Owns the Data? Open Data for Healthcare

Patty Kostkova^{1*}, Helen Brewer², Simon de Lusignan³, Edward Fottrell⁴, Ben Goldacre⁵, Graham Hart⁴, Phil Koczan⁶, Peter Knight⁷, Corinne Marsoller⁸, Rachel A. McKendry⁹, Emma Ross¹⁰, Angela Sasse¹, Ralph Sullivan¹¹, Sarah Chaytor¹², Olivia Stevenson¹², Raquel Velho¹³ and John Tooke¹⁴

Research on large shared medical datasets and data-driven research are gaining fast momentum and provide major opportunities for improving health systems as well as individual care. Such open data can shed light on the causes of disease and effects of treatment, including adverse reactions side-effects of treatments, while also facilitating analyses tailored to an individual's characteristics, known as personalized or "stratified medicine." Developments, such as crowdsourcing, participatory surveillance, and individuals pledging to become "data donors" and the "quantified self" movement (where citizens share data through mobile device-connected technologies), have great potential to contribute to our knowledge of disease, improving diagnostics, and delivery of healthcare and treatment. There is not only a great potential but also major concerns over privacy, confidentiality, and control of data about individuals once it is shared. Issues, such as user trust, data privacy, transparency over the control of data ownership, and the implications of data analytics for personal privacy with potentially intrusive inferences, are becoming increasingly scrutinized at national and international levels. This can be seen in the recent backlash over the proposed implementation of care.data, which enables individuals'

devices, the IT industry and MedTech giants are pursuing new projects without clear public and policy discussion about ownership and responsibility for user-generated data. In the absence of transparent regulation, this paper addresses the opportunities of Big Data in healthcare together with issues of responsibility and accountability. It also aims to pave the way for public policy to support a balanced agenda that safeguards personal information while enabling the use of data to improve public health.

information while enabling the use of data to improve public health.

Numerous scientists have pointed out the irony that right at the historical moment when we have the technologies to permit worldwide availability and distributed process of scientific data, broadening collaboration and accelerating the pace and depth of discovery ... **we** are busy locking up that data and preventing the use of correspondingly advanced technologies on knowledge.



John Wilbanks, VP Science, Creative Commons