Istituto Superiore di Sanità

National Institute of Health in Italy

Research for public health

The Istituto Superiore di Sanità (ISS, the Italian National Institute of Health) is the main Italian research institute in the biomedical and public health field. It is the technical and scientific body of the Italian National Health Service.

MISSION

Promotion and protection of national and international public health through research, surveillance, regulation, control, prevention, communication, counselling and training.

ISS MAIN FIELDS OF ACTIVITY

- Cardiovascular, endocrine-metabolic diseases and ageing
- Environment and health
- Food safety, nutrition and veterinary public health
- Infectious diseases
- Neuroscience
- Oncology and molecular medicine
- Addiction and doping
- Animal research and welfare
- Chemicals, cosmetics and consumer protection
- Clinical excellence, healthcare quality and safety
- Control and evaluation of medicines
- Disease prevention and health promotion
- Drug research and evaluation
- Global health
- Health technology assessment
- HIV/AIDS research
- Innovative technologies in public health
- Radiation protection and computational physics
- Rare diseases
- Telemedicine and new healthcare technologies
- Blood Transplant

The available information assets

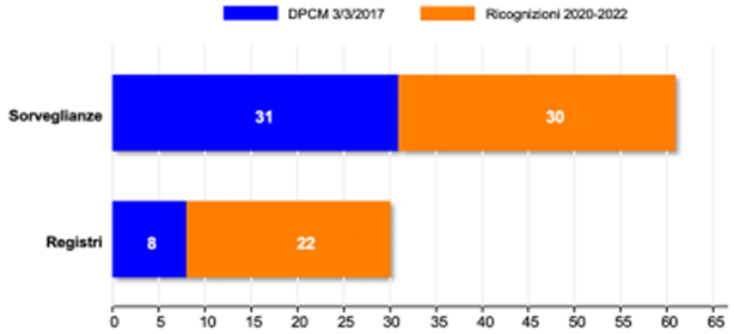


Figura 10. Distribuzione complessiva di Sorveglianze/Registri affidati all'ISS dal DPCM 3/3/2017 o identificati nelle indagini 2020 e 2022 (attivi al 31 dicembre 2022)

STATISTICS Service

Director: Giada Minelli
giada.minelli@iss.it

- **Represents ISS in the National Statistical System, carrying out all the activities established by law.**
- **Carries out statistical and epidemiological analyses of health and demographic data from current flows and international collaborations.**
- **Performs health statistics activities.**

ISS STATISTICS SERVICE DATABASE

The STATISTICS SERVICE has access to Public Health data from various Institutions. These data are then organized in databases.

Our main databases are:

MORTALITY DATA

source: Istat
individual data including causes of death
available from 1980

HOSPITALIZATION DATA

source: Italian Ministry of Health
individual data including discharge diagnosis
available from 2001

Internship Day 2024
Sapienza – Department of Statistical Sciences
October 18th

Pharmacoepidemiology and Pharmacosurveillance Unit

National Center for Drug Research and Evaluation
Italian National Institute of Health
Rome, Italy

Marco Massari, Flavia Mayer, Stefania Spila Alegiani

ISTITUTO SUPERIORE DI SANITÀ
CENTRO NAZIONALE
RICERCA E VALUTAZIONE PRECLINICA
E CLINICA DEI FARMACI

Pharmacoepidemiology

Study of interactions between drugs and human populations, investigating, in real conditions of life (**Real World**), benefits, risks and use of drugs.

Pharmacoepidemiology applies to drugs the methods of Pharmacology, Epidemiology, Statistics and Data Science.

We mainly collaborate with Italian Regions and with:

AIFA AGENZIA ITALIANA DEL FARMACO

EUROPEAN MEDICINES AGENCY SCIENCE MEDICINES HEALTH

AIE farmacoepi

UNIVERSITÀ DEGLI STUDI DI VERONA

DIEP Lazio
Dipartimento di Epidemiologia del Servizio Sanitario Regionale Regione Lazio

Real World Data at Regional Level

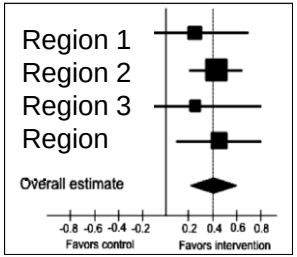
- Population Registry
- Pharmacy Claims
- Hospitalization Discharges
- Emergency Care Visits
- Copayment Exemptions
- Outpatients Specialist Registry
- Certificate of childbirth assistance (CEDAP)
- Other local Registries:
  - Covid-19 Surveillance Registry
  - Vaccination Registry
  - etc.

Regional-level Record Linkage

Strategies to conduct a pharmacoepidemiological study

SHARING OF RAW DATA
Pooled raw data anaysis

LOCAL ANALYSES
No raw data sharing
Sharing only results



COMMON DATA MODEL (CDM)
Distributed Analysis

TheShinISS

***Source**: The European Network of Centres for Pharmacoepidemiology and Pharmacovigilance (ENCePP). Guide on methodological standards in Pharmacoepidemiology.*
http://www.encepp.eu/standards_and_guidances/methodologicalGuide.shtml

TheShinISS

- A tool (R/Shiny-based) to conduct distributed analyses in pharmacoepidemiology
- A large scale distributed database network for postmarketing surveillance of drugs and vaccines
- Some key research questions during COVID:

What is the survival rate of COVID-19 hospitalized patients, stratified by age, sex, and risk factors?

Clin Epidemiol 2020;12:1337–46

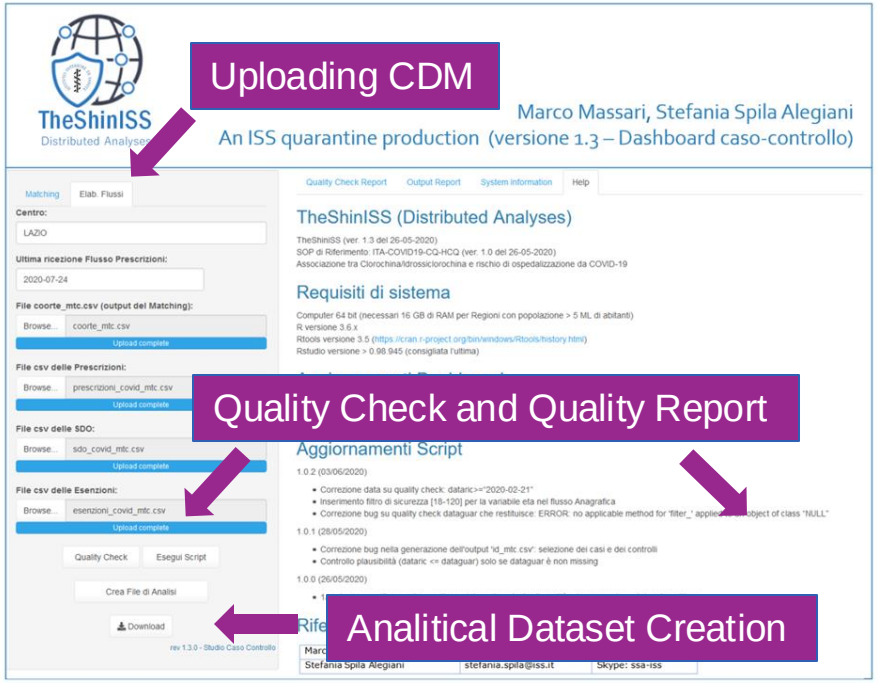
Can hydroxychloroquine modify the prognosis of COVID-19?

Rheumatology (Oxford). 2021;60(SI):SI25–SI36

Are COVID-19 mRNA vaccines associated with myocarditis/pericarditis in the population younger than 40 years?

Plos Medicine 2022, 19(7): e1004056

studies
n. ~ 40.000 – 12.5 ML individuals

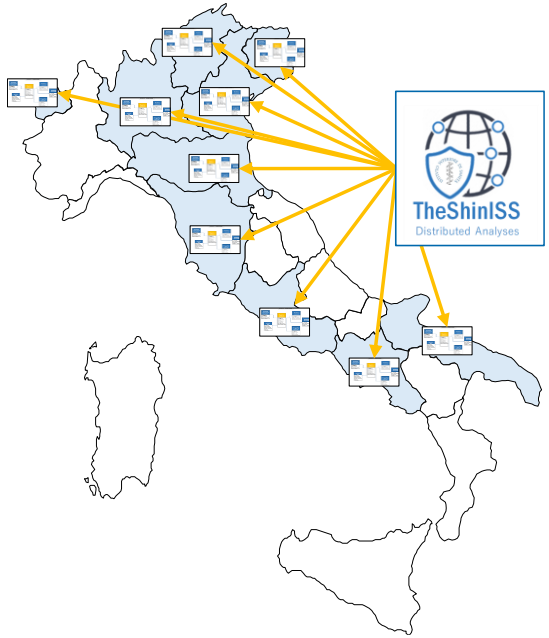


Massari M et. al. TheShinISS: un applicativo open-source per la conduzione di analisi distribuite in studi di farmacoepidemiologia di tipo multi-database. Boll Epidemiol Naz 2020; 1(2):39-45. DOI: https://doi.org/10.53225/BEN_006

Distributed Analysis with TheShinISS

TheShinISS is distributed and installed in all regions.

Raw data, structured according to the Common Data Model, are LOCALLY analysed by TheShinISS ⇨ DISTRIBUTED ANALYSIS



Tasks of TheShinISS

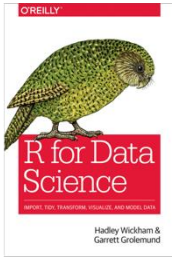
-Quality Checks on local databases
-Identification of the study population (es. by matching)
-Record Linkage beetween databases
-Filters, aggregation and re-coding
-Covariates, expositions and outcomes construction
-Creation of the Anonymized Analytical Dataset
-Aggregated Results (when required)

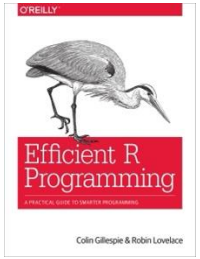
Data Science Skills

- Collaborative Development with GitHub

- Reproducible Analysis (R Markdown, Quarto)

- Efficient Programming in R





- Tidyverse and Shiny

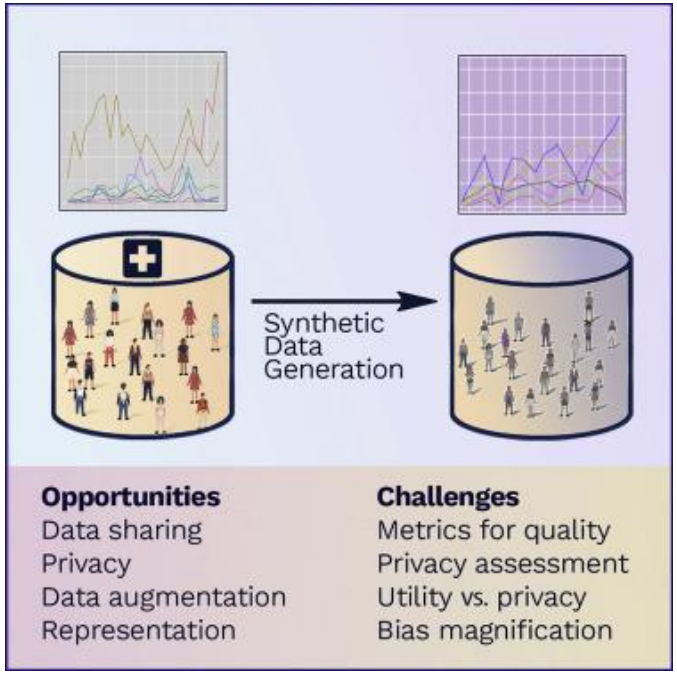
- Querying Servers via R

- Defensive Programming (check functions to ensure Code Integrity)

Synthetic data

- “Synthetic data generation is the process of using machine learning methods to train a model that **captures the patterns in a real dataset**.

- The synthetic **data does not have a one-to-one mapping to the original data or to real patients**, and therefore has the potential of **privacy preserving properties**.”



Rajotte iScience. 2022 Nov 18; 25(11): 105331.

Generative Adversarial Networks (GANs)

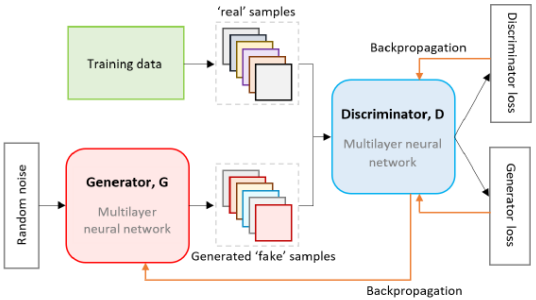
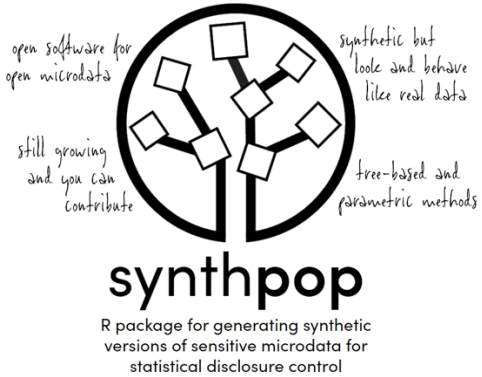


Figure 1: Example of GAN Architecture

Little C, Elliot M, Allmendinger R, et al. Generative Adversarial Networks for Synthetic Data Generation: A Comparative Study. arXiv **2021**, arXiv:2112.01925v1.

GANs, introduced by Ian Goodfellow in 2014, are a type of artificial intelligence model composed of two neural networks, a generator and a discriminator, that work in competition with each other.

Synthpop



https://www.synthpop.org.uk/links.html

Synthpop is a free tool written in R. For each variable in the real data (RD), you can specify whether to use parametric methods or non-parametric methods, primarily CART

Grazie per l’attenzione !

marco.massari@iss.it
flavia.mayer@iss.it
stefania.spila@iss.it