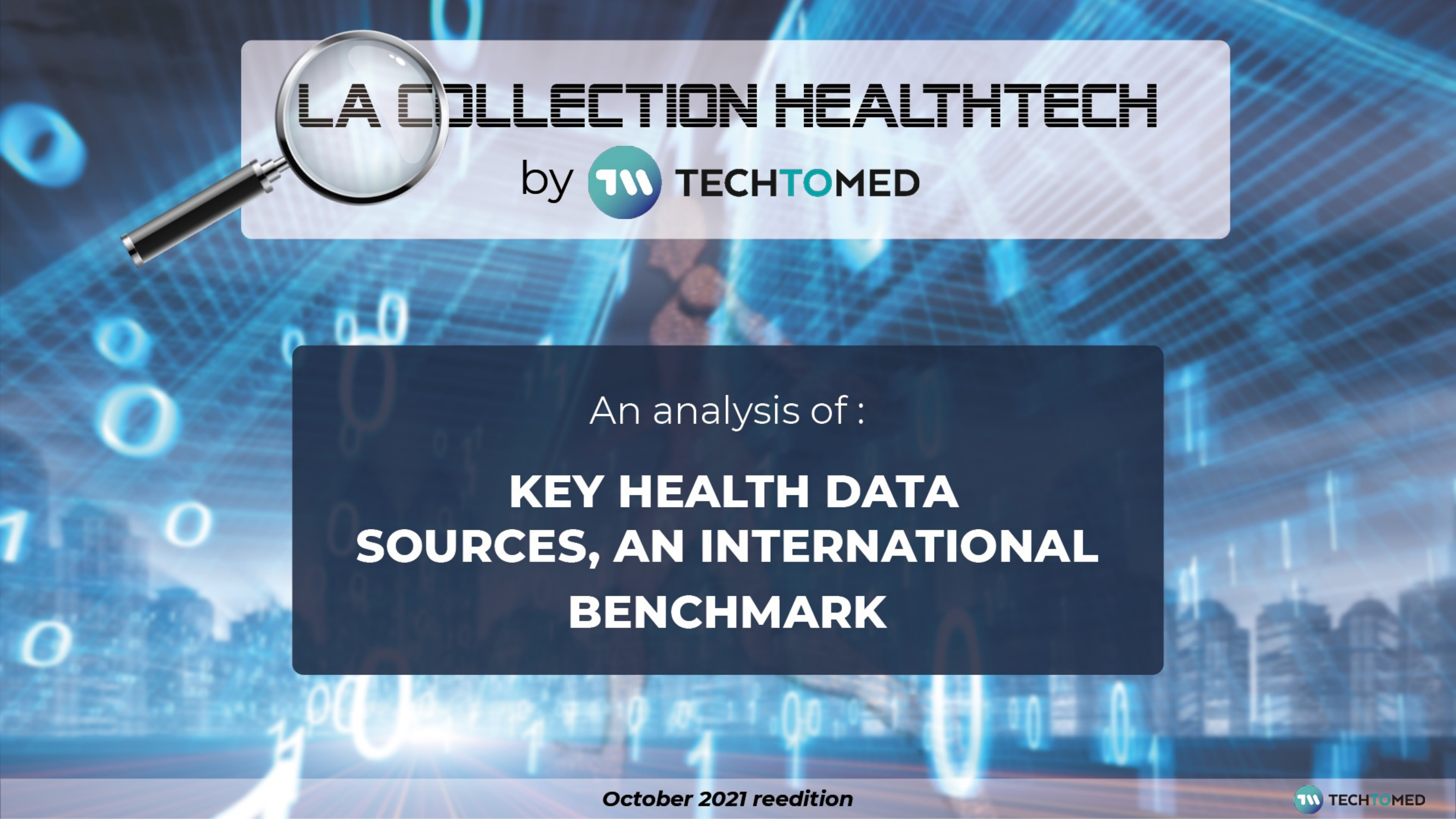


LA COLLECTION HEALTHTECH

by TECHTOMED

An analysis of :

KEY HEALTH DATA SOURCES, AN INTERNATIONAL BENCHMARK

October 2021 reedition

TECHTOMED

CONTENTS

CREDITS

About the authors

Techtomed SAS is Paris based company, acting as a consulting organization and dedicated to HealthTech field.
Techtomed team is composed with healthtech consultants and an experts network.

Franck Le Meur is the Founder and CEO of Techtomed SAS.

Anca Petre is the co-founder and CEO of 23 Consulting, a Paris-based firm specialized in tracking the latest technological innovation with high potential for patients. In her work, Anca trains and assists major players in this sector on an international scale, including patient advocacy groups, pharmaceutical corporations, healthcare providers and international NGOs.

Legal Notice

All contents of Service are the exclusive intellectual property of Techtomed SAS including trademarks, copyrights or any other form of intellectual property. Terms of a license provided by Company are governed by the Purchase and Licensing Agreement below.

Service licenses are designed to regulate readership and distribution rights.

1. The purchase of a single user license grants access to a specific report for one person only and must not be shared with other employees within the same company. Option A is available for industries and large companies (more than 50 employees). Option B concerns only start-up companies, SMEs (less than 50 employees) and public institutions.
2. For an extended license (multiple users, within the same department or company) or any other related specific requirement, please contact us at contact@techtomed.com

The report will be delivered to you online/by email in a secured PDF format within 24 hours of receipt of the completed and signed purchase order form.

Refund Policy : Please read all the available information about any report before placing your order. Due to the confidential nature of the information contained in our reports, cancellation of orders cannot be accepted after the report has been delivered.

METHODOLOGY

The methodology is based on desk research

Scientific publications
Specialized journals
Websites from institutions, corporations and startups
Studies and institutional recommendations

1 Identify the perimeter of health data and key international trends

1. Study the definition of “health data” to define a perimeter for the study.
2. Identify and synthesize the different types and sources of health data, with their strengths and weaknesses.
3. Identify and synthesize 10 trends in health data to anticipate the evolution in key health data sources globally.

2 Map national and international health databases

1. Select 4 types of key data sources that present the following criteria:
    - International experience with these data sources ;
    - Large international spread ;
    - Known and used by key healthcare stakeholders ;
    - Of interest given the previously identified trends

4 types of data sources were selected in this study :

- Public national databases;
- National Electronic Health Records systems;
- International cohorts;
- Real-World Data databases.

2. For each type of data source, identify the countries with the following criteria :
    - High volume of information available;
    - Geographical representativeness;
    - Technological maturity.
3. Compare countries for each type of data source.

3 Identify key health data use cases

1. Identify the initiatives and projects making use of large volumes of health data.
2. Select the initiatives that rely of the data sources previously identified and in accordance with the trends presented in part 1.
3. Identify the key stakeholders involved in these use cases, their strengths and their weakness.
4. Assess the spread of these use cases at an international level by mapping initiatives in key countries.

EXECUTIVE SUMMARY

Context and objectives

”Health data” is a generic term that refers to a wide range of data types that have a common feature: they inform about the health status of an individual or a population. Such data can be found in multiple sources scattered around the healthcare system.

The variety of health data types and sources, combined with the complexity in identifying and evaluating them, motivated the creation of this report by TechToMed.

“International benchmark of key health data sources” is the first report to map and assess the most important sources of health data in the world.

The objectives of this study are:
- Map and assess the different sources of health data worldwide ;
- Identify key use cases and trends based on these data sources ;
- Propose recommendations to healthcare stakeholders in their health data strategies.

What you will find in this study

- A complete overview of all data types and sources with assessments of their specificity, their consistency, their comprehensiveness, they level of standardization, their interoperability and their accessibility.
- Top 10 trends in health data worldwide illustrated by facts and figures.
- An in-depth description, analysis and evaluation of 4 main categories of health data sources: national databases, electronic health records, cohorts and real-world databases in an international scale.
- An overview of international data protection and security frameworks including international interoperability standards.
- 5 of the most important use cases in health data: the European digital single market, public health monitoring with a focus on Covid-19, personalized medicine, oncology research and pharmaceutical development.

GLOSSARY

- **GDPR :** General Data Protection Regulation
- **EHR :** Electronic Health Record
- **IoT :** Internet of Things
- **OECD :** the Organisation for Economic Co-operation and Development
- **RCT :** Randomized controlled trials
- **RWD :** Real World Data
- **DTx :** Digital therapeutics
- **GDP :** Gross Domestic Product
- **PROM :** Patient reported outcome measures
- **PREM :** Patient reported experience measures
- **SNDS :** Système national de données de santé
- **NHS :** National health system
- **HIRA :** Health insurance review & assessment service
- **HSCIC :** Health and social care information center
- **NCHS :** National center for health statistics
- **CIHI :** Canadian Institute for Health Information
- **ERB :** Ethics review board
- **DMP :** Dossier medical partagé
- **CNAM :** Caisse nationale d’assurance maladie ID : Identification
- **IHCC :** International hundredK+ cohorts consortium
- **PEACE :** Patient-centered evaluative assessment of cardiac events
- **CPS :** Cancer prevention study
- **AMORIS :** Apolipoprotein mortality risk study
- **EPIC :** European prospective investigation into cancer, chronic diseases, nutrition and lifestyle
- **CHOP :** Children’s hospital of Philadelphia GS : Generation study
- **SNIIRAM :** Système national d’information inter-régimes de l’Assurance maladie
- **PMSI :** Programme de médicalisation des systèmes d’information
- **CMS :** Center for Medicare & Medicaid services
- **CPRD :** Clinical practice research datalink
- **THIN :** The health improvement network
- **HES :** Hospital episode statistics
- **PCORnet :** The national patient-centered clinical research network
- **HIRD :** Healthcore integrated research database
- **DP :** Data partners
- **SOC :** Sentinel operations center
- **INDS :** Institut national des données de santé
- **CEREES :** Comité d’expertise pour les recherches, les études et les évaluations dans le domaine de la santé
- **CNIL :** Commission nationale de l’informatique et des libertés
- **OPCRD :** Optimum patient care research database

GLOSSARY

- **ACP :** Accelerated cure project
- **HIPAA :** Health insurance portability and accountability act
- **PHI :** Protected health information
- **IHI :** Identifiable health information
- **OMOP :** Observational medical outcomes partnership
- **HL7 :** Health lever 7
- **FHIR :** Fast healthcare interoperability resources
- **DICOM :** Digital imaging and communications in medicine
- **ICPC :** International classification of primary care
- **ICD-10 :** International statistical classification of diseases and related health problems, tenth revision
- **SNOMED CT :** Systematized nomenclature of human and veterinary medicine clinical terms
- **LOINC :** Logical observation identifiers names & codes
- **SOAP :** Subjective, objective, assessment, plan standard
- **CDA :** Clinical document architecture
- **RIM :** Reference information model
- **CEN :** Comité européen de normalization
- **ISO :** International standards organization
- **IHE :** Integrating the healthcare enterprise
- **DIN :** Drug identification number
- **OPCS :** Oligodendrocyte progenitor cells
- **DSM-IV :** Diagnostic and statistical manual of mental disorders
- **UCUM :** The unified code for units measure
- **NABM :** Nomenclature des actes de biologie médicale
- **HCPCS :** Healthcare common procedure coding system
- **IMI :** Innovative medicines initiative
- **EFPIA :** The European federation of pharmaceutical industries and associations
- **BDV :** Big data value association
- **EHDEN :** European health data & evidence network
- **EMIF :** European medical information framework
- **EHR4CR :** Electronic health records for clinical trials
- **PEPP-PT :** Pan-European privacy-preserving proximity tracing
- **INRIA :** Institut National de recherche en informatique et en automatique
- **ICGC :** International cancer genome consortium

SUMMARY

- Perimeter and trends in health data
- Global health data sources map
- International data protection and security frameworks
- Health data use cases

Definition of health data

In Europe, the GDPR gives the first comprehensive definition of health data

The GDPR defines health data

Health data is data concerning health' means personal data related to the physical or mental health of a natural person, including the provision of health care services, which reveal information about his or her health status.

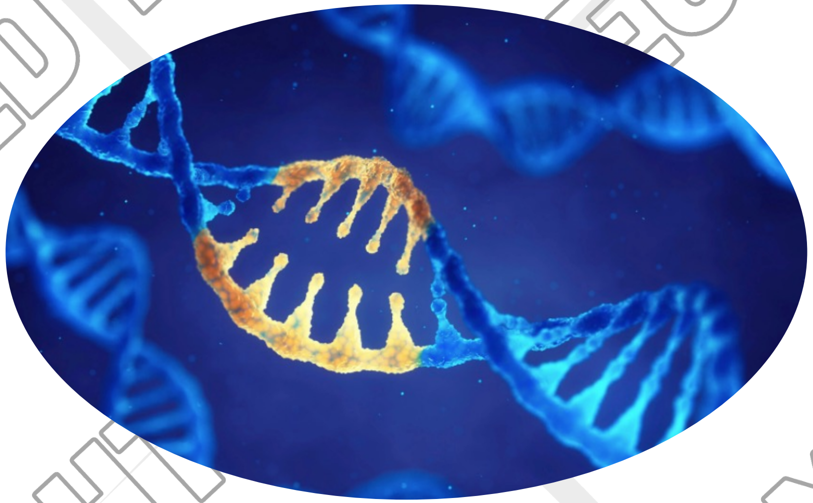
Health data can be about an individual's past, current or future health status. It not only covers specific details of medical conditions, tests or treatment, but includes any related **data which reveals anything about the state of someone's health**.

This includes:

- Data collected during the registration for, or the provision of, health care;
- A number or symbol assigned to an individual to uniquely identify him for health purposes;
- Data from the testing or examination of a body part or bodily substance;
- Genetic data and biological samples;
- Information on a disease, disability, disease risk, medical history, clinical treatment;
- Data on the physiological or biomedical state of the data subject independent of its source, for example from a physician or other health professional, a hospital, a medical device or an in vitro diagnostic test.

Official Journal of the European Union, Regulation (EU) 2016/679
Data Resources in the Health Sciences, Health Sciences Library, University of Washington

Genetic data is defined separately



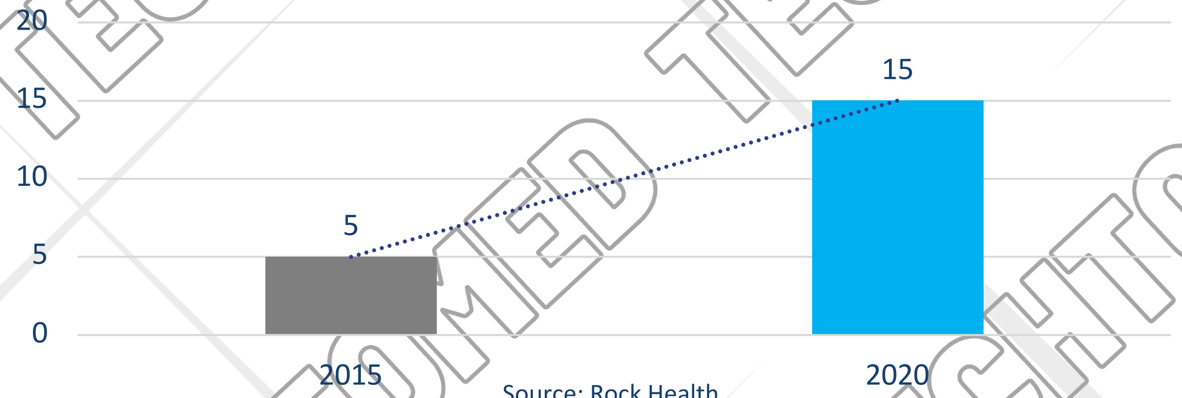
The GDPR defines genetic data as "personal data relating to the inherited or acquired genetic characteristics of a natural person which result from the analysis of a biological sample from the natural person in question, in particular chromosomal, deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) analysis, or from the analysis of another element enabling equivalent information to be obtained."

Genetic data is a special category of health data as it is highly sensitive as it could create significant risks for an individual's privacy rights and fundamental freedom.

Trend #7: The Internet of Medical Things will become mainstream, especially for physical activity and cardiac monitoring

The global IoT spending in healthcare is expected to multiply by 3 in 5 years, fueled by skyrocketing adoption rates from the general public

Health IoT spendings by consumers in $bn



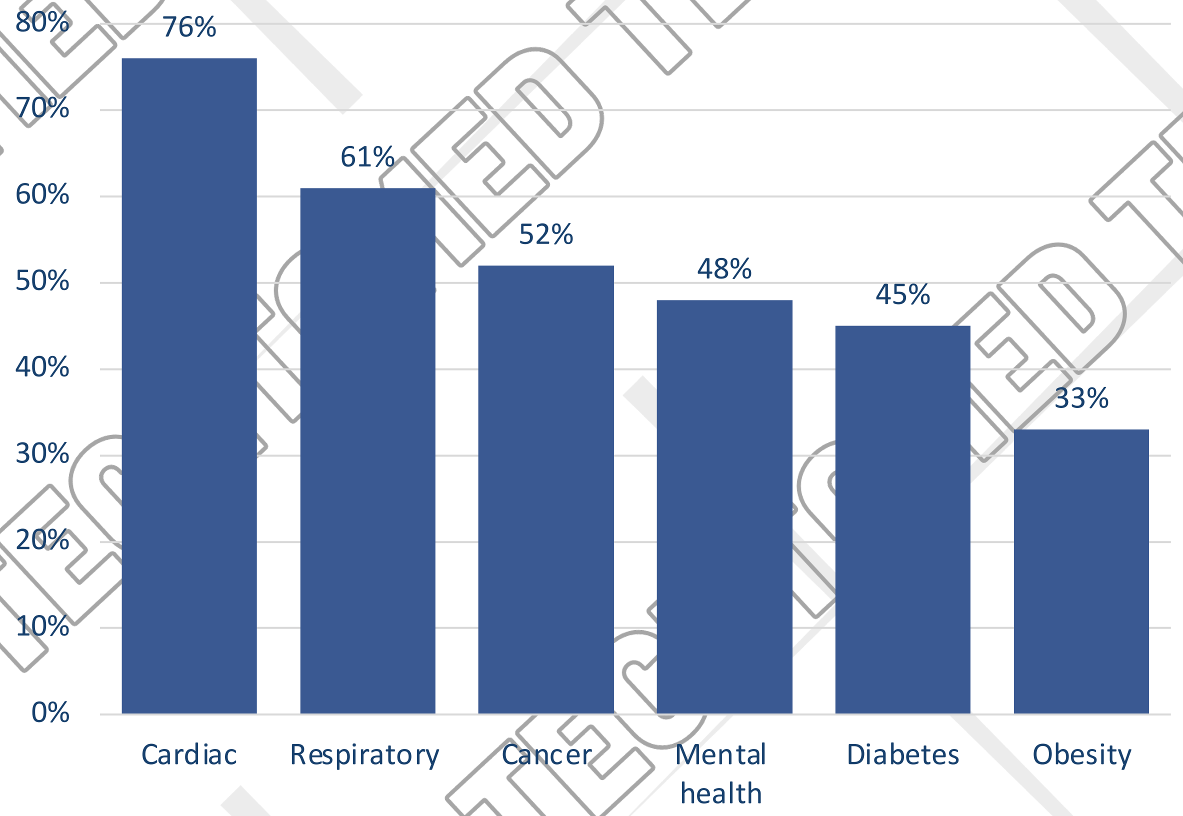
Source: Rock Health

Reasons for wearable use

According to : Rock Health

Connected medical devices and wearables allow patients to monitor their chronic conditions from home.

Conditions tracked by health providers through remote patient monitoring

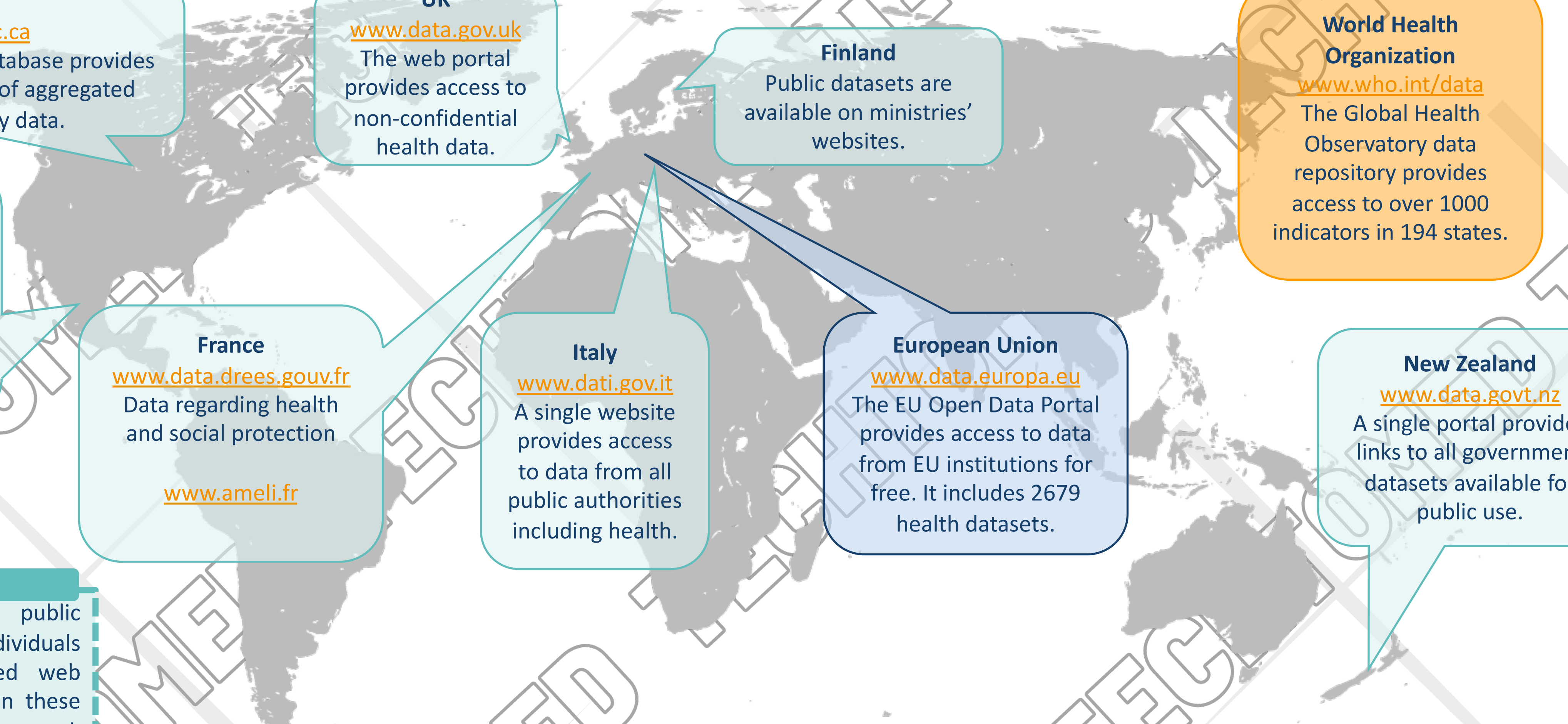


According to : Rock Health

SUMMARY

- Perimeter and trends in health data
- Global health data sources map
- International data protection and security frameworks
- Health data use cases

In addition to national databases, countries all over the world develop open data initiatives



Canada
www.data.gc.ca
The government-wide database provides access to a wide range of aggregated health and safety data.

UK
www.data.gov.uk
The web portal provides access to non-confidential health data.

Finland
Public datasets are available on ministries' websites.

World Health Organization
www.who.int/data
The Global Health Observatory data repository provides access to over 1000 indicators in 194 states.

US
www.healthdata.gov
The initiative includes aggregated health data and is managed by the Department of Health and Human Services.

France
www.data.drees.gouv.fr
Data regarding health and social protection

www.ameli.fr

Italy
www.dati.gov.it
A single website provides access to data from all public authorities including health.

European Union
www.data.europa.eu
The EU Open Data Portal provides access to data from EU institutions for free. It includes 2679 health datasets.

New Zealand
www.data.govt.nz
A single portal provides links to all government datasets available for public use.

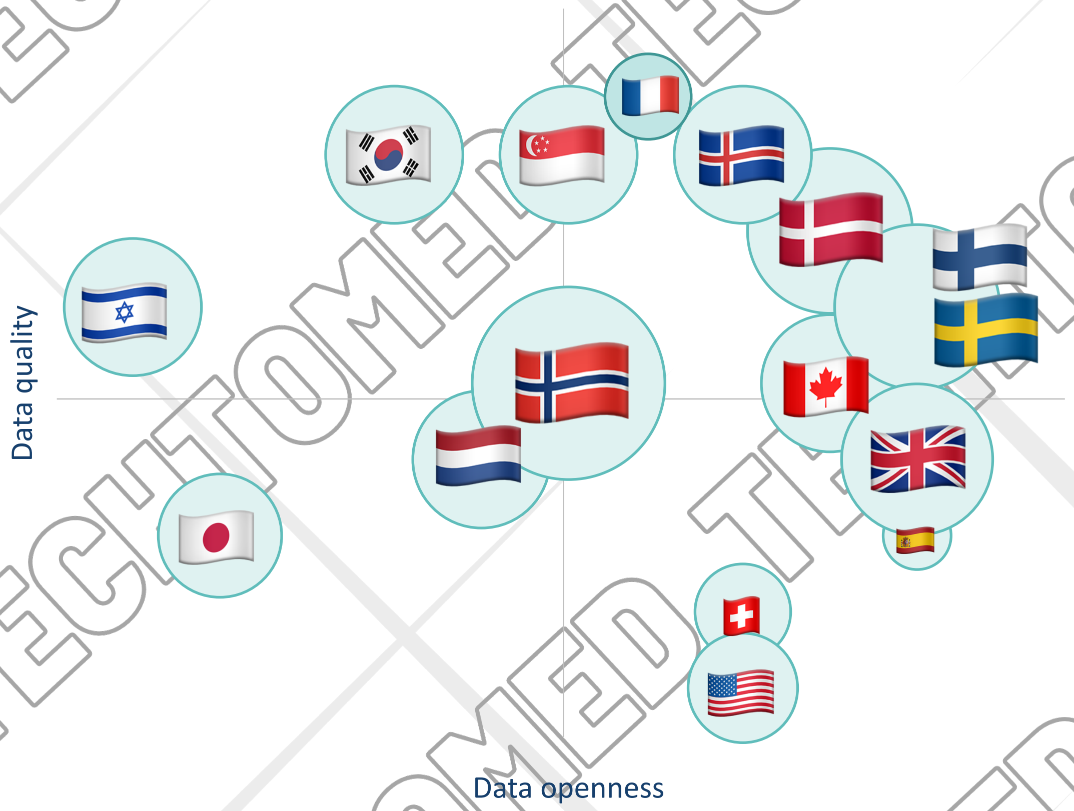
Definition

Open data initiatives are public databases accessible by all individuals usually through a dedicated web platform. All data available in these databases can be freely accessed, used and shared, independently of the scope.

Assessment of national databases in the USA, the UK, Canada, Finland and Korea (10/10)

Summary (2/2)

Global country score for national databases



The width of the bubbles represents the availability of key datasets in national databases

Where does France stand?

France's main national database system is the SNDS (Système National de Données de Santé) which includes hospital discharge data, claims and mortality registers.

The French system has 3 main strengths:

- A large sample size: the whole French population
- Large, comprehensive and detailed information, without any exclusion
- No loss of follow-up, except if long stay abroad

However, the national database does not include the following data:

- Over-the-counter drugs
- Results of clinical examinations
- Results of laboratory tests
- Lifestyle data, PROM and PREM
- Long-term care data

The SNDS is opened to any individual or organization, private or public, commercial or non-profit granted access through an application process.

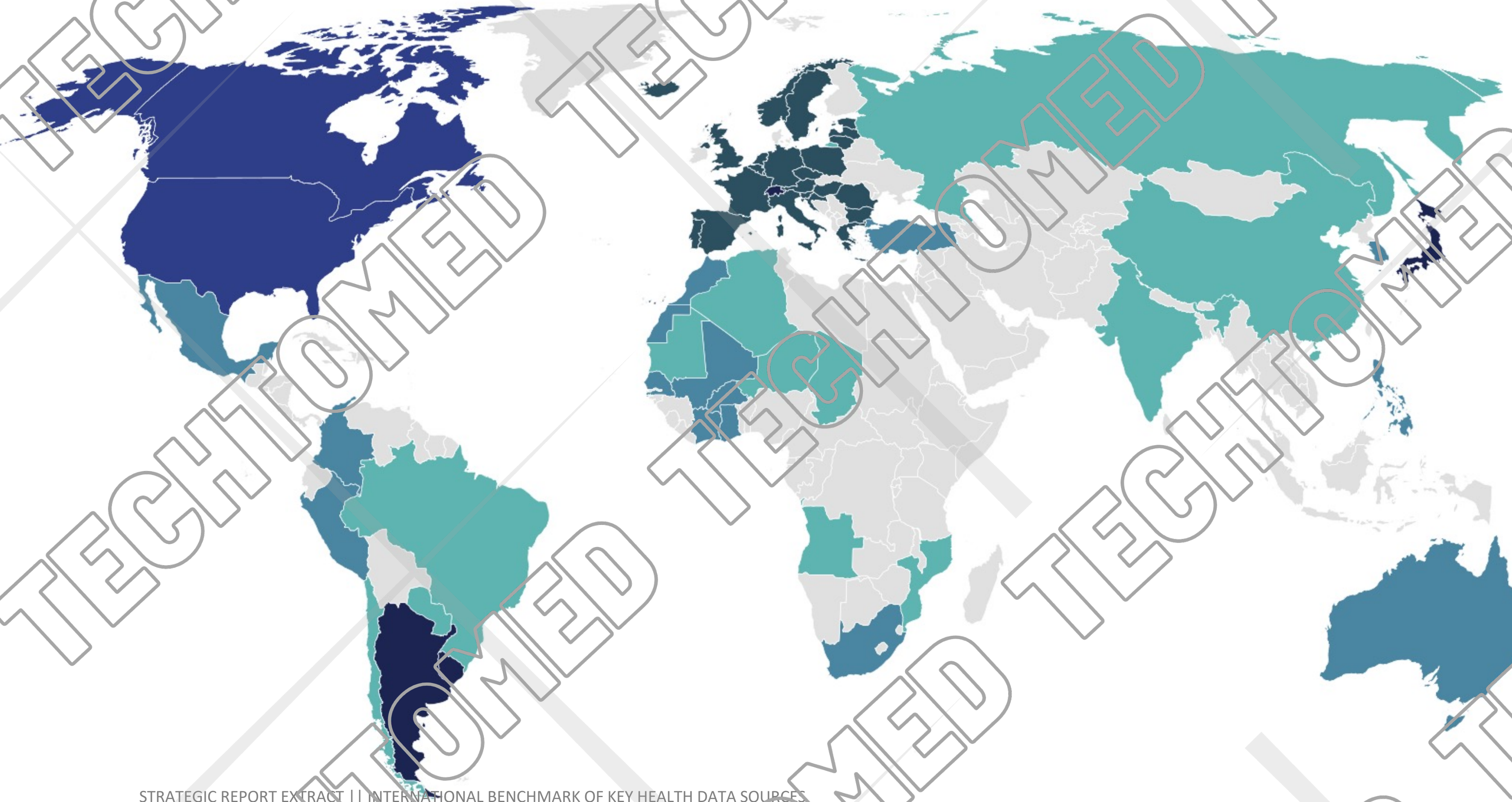
Regarding the quality of the data, there are irregularities in the coding and data is not consistently collected.

According to : AMELI, La Cour des Comptes

SUMMARY

GDPR is considered as the most constraining regulation

Level of data protection across the world in comparison to the European requirements



The French data protection authority (CNIL) mapped data protection policies in various countries and concluded on their compatibility with the GDPR.

Switzerland, Argentina and Japan are the only country with a perfectly adequate data protection system when compared with the GDPR.

Measuring the degree of individual control over health data

Different aspects of health data over which patient have control

	🇦🇺	🇦🇹	🇨🇦	🇨🇳	🇩🇰	🇪🇪	🇫🇮	🇮🇸	🇮🇱	🇯🇵	🇳🇱	🇳🇴	🇸🇬	🇪🇸	🇸🇪	🇨🇭	🇬🇧	🇺🇸
Allows individuals electronic access to their own health-related data when held in an EHR	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟥	🟩	🟥	🟥	🟩	🟩	🟩	🟥	🟩	🟩
Allows individuals to demand their own health-related data be corrected when held in an EHR if it is k0wn to be inaccurate	🟩	🟩	🟩	🟥	🟩	🟩	🟩	🟩	🟥	⬜	🟩	🟥	🟩	🟩	🟩	🟥	🟥	🟩
Allows individuals to demand the deletion of health-related data from their EHR	🟩	🟩	🟥	🟥	🟩	🟩	🟥	🟩	🟥	⬜	🟩	🟥	🟥	🟩	🟩	🟥	🟥	🟥
Allows individuals to specify which health-related data from their EHR can be shared with health professionals of their choice	🟩	🟩	🟥	⬜	🟩	🟩	🟩	🟩	🟩	⬜	🟩	🟥	🟥	🟩	🟩	🟥	🟩	🟥
Score	4	4	2	0	3	4	3	4	1	1	3	0	2	4	4	0	1	2

🟩 Control 🟥 No control ⬜ Unknown

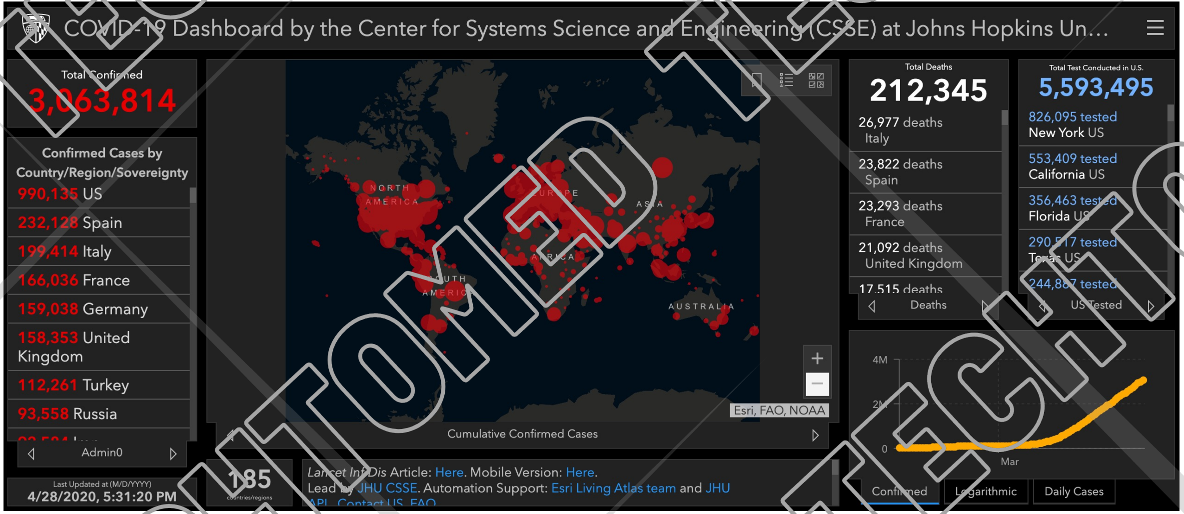
World Health Organization

SUMMARY

- Perimeter and trends in health data
- Global health data sources map
- International data protection and security frameworks
- Health data use cases

Case study: using health data to monitor public health and performance during the COVID-19 pandemic

Many organizations built dashboard from open source data to visualize key indicators of the epidemic


COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins Un…

Total Confirmed
3,063,814

Confirmed Cases by Country/Region/Sovereignty
990,135 US
232,128 Spain
199,414 Italy
166,036 France
159,038 Germany
158,353 United Kingdom
112,261 Turkey
93,558 Russia

Last Updated at (M/D/YYYY)
4/28/2020, 5:31:20 PM

185
countries/regions

Lancet Inf Dis Article: Here. Mobile Version: Here.
Lead by JHU CSSE. Automation Support: Esri Living Atlas team and JHU

Total Deaths
212,345
26,977 deaths Italy
23,822 deaths Spain
23,293 deaths France
21,092 deaths United Kingdom

Total Test Conducted in U.S.
5,593,495
826,095 tested New York US
553,409 tested California US
356,463 tested Florida US
290,517 tested Texas US

Depending on the organization and on the available data sources, information on dashboards included:

- Track of confirmed new cases, hospitalization, deaths by country of healthcare provider
- Number of tests and rate of positive or negative tests
- News reports
- Scientific publications
- Tweets
- Clinical trials
- Patient transfers in-country or between countries

A large number of health stakeholders offered public access to their datasets for research purposes in an unprecedented wave of national international collaborations.

🇫🇷

Hospitals in the Paris area (AP-HP) created a COVID-19 task force to analyze and track the epidemic through the data collected in the hospitals' data warehouses. This led to the creation of the EDS-Covid-19 database, specifically designed for the epidemic.

The national data platform, Health Data Hub, and the national payer, CNAM, were allowed to massively collect data related to COVID-19 in order to facilitate research.

🇪🇺

The European Commission established the ERAvsCORONA action plan that coordinates efforts and supports research against COVID-19. One of the measures of the plan is the development of a data storage and sharing platform for COVID-19-related data.

🇺🇸

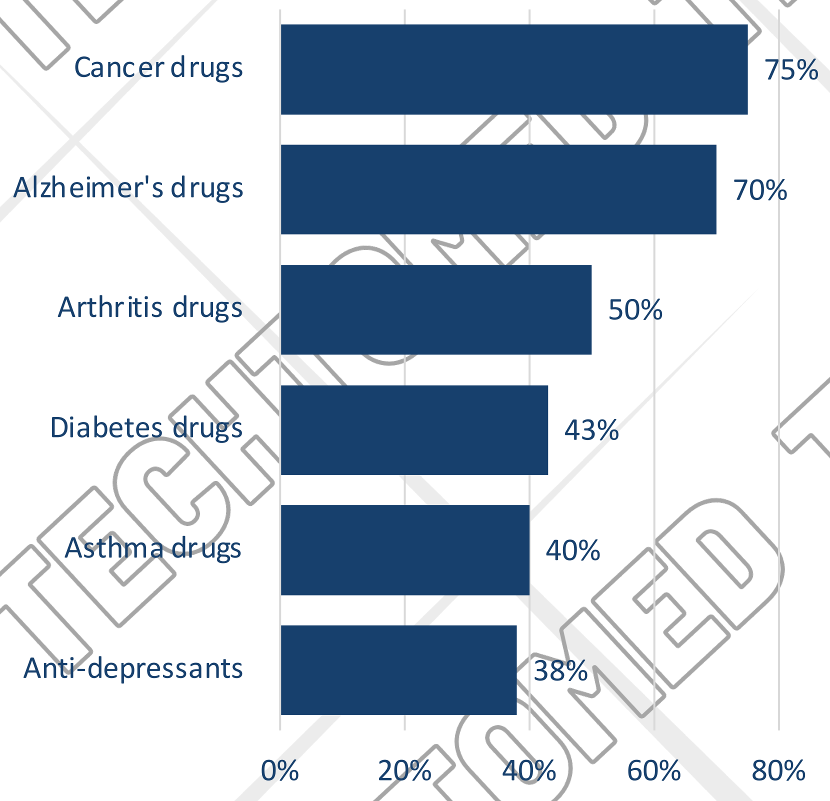
28 US firms formed a collaboration on an EHR-fueled database accessible to doctors. They will be able to verify drug effectiveness, identify pre-existing conditions that can lead to ventilator support and evaluate the impact of lockdown measures.

Cleveland Clinic and SAS have co-created predictive models to help hospitals pan and allocate resources in the context of the epidemic

Personalized medicine uses data to increases the chances of positive outcomes

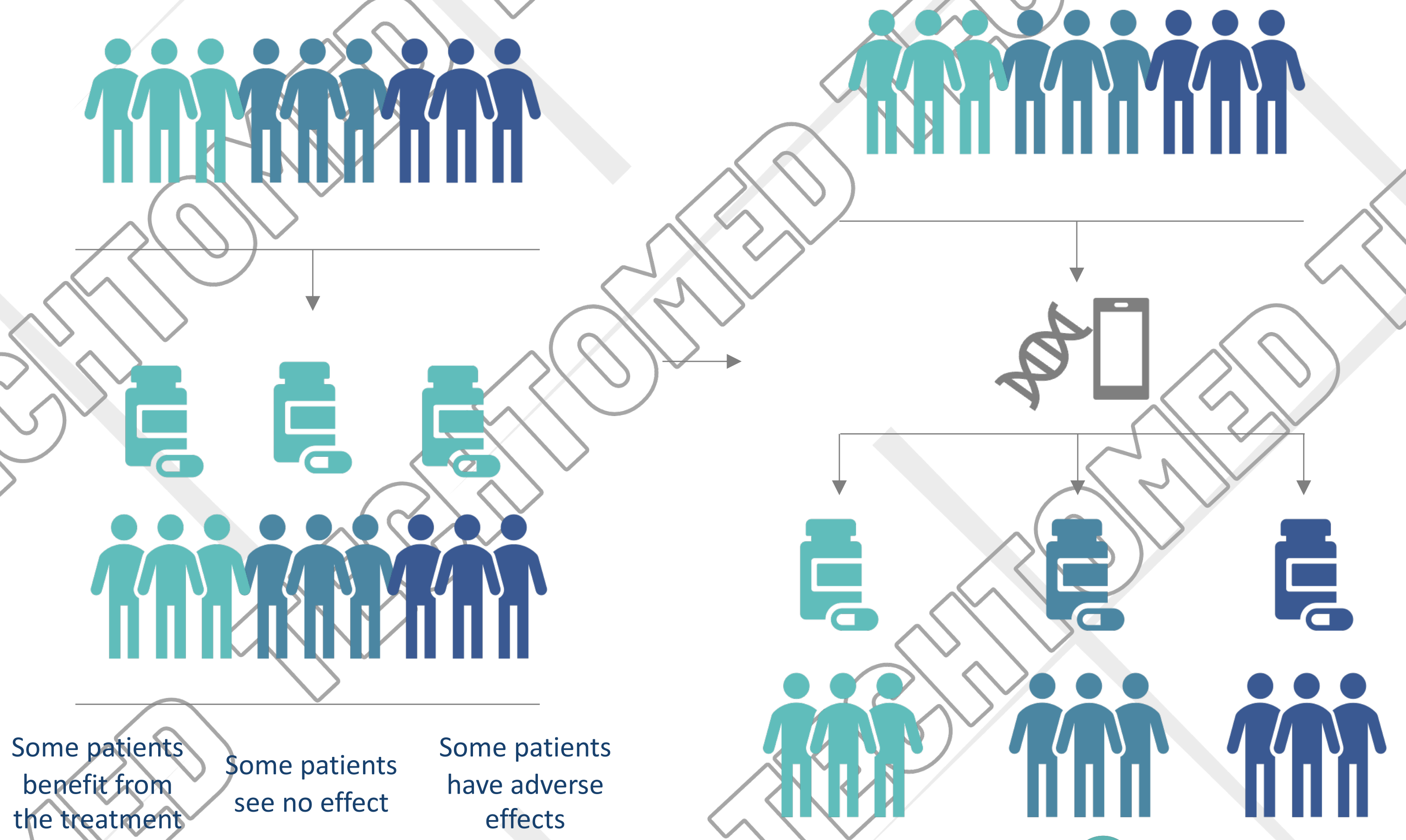
With personalized medicine, access to patient data allows healthcare providers to design individualized treatments that improve outcomes.

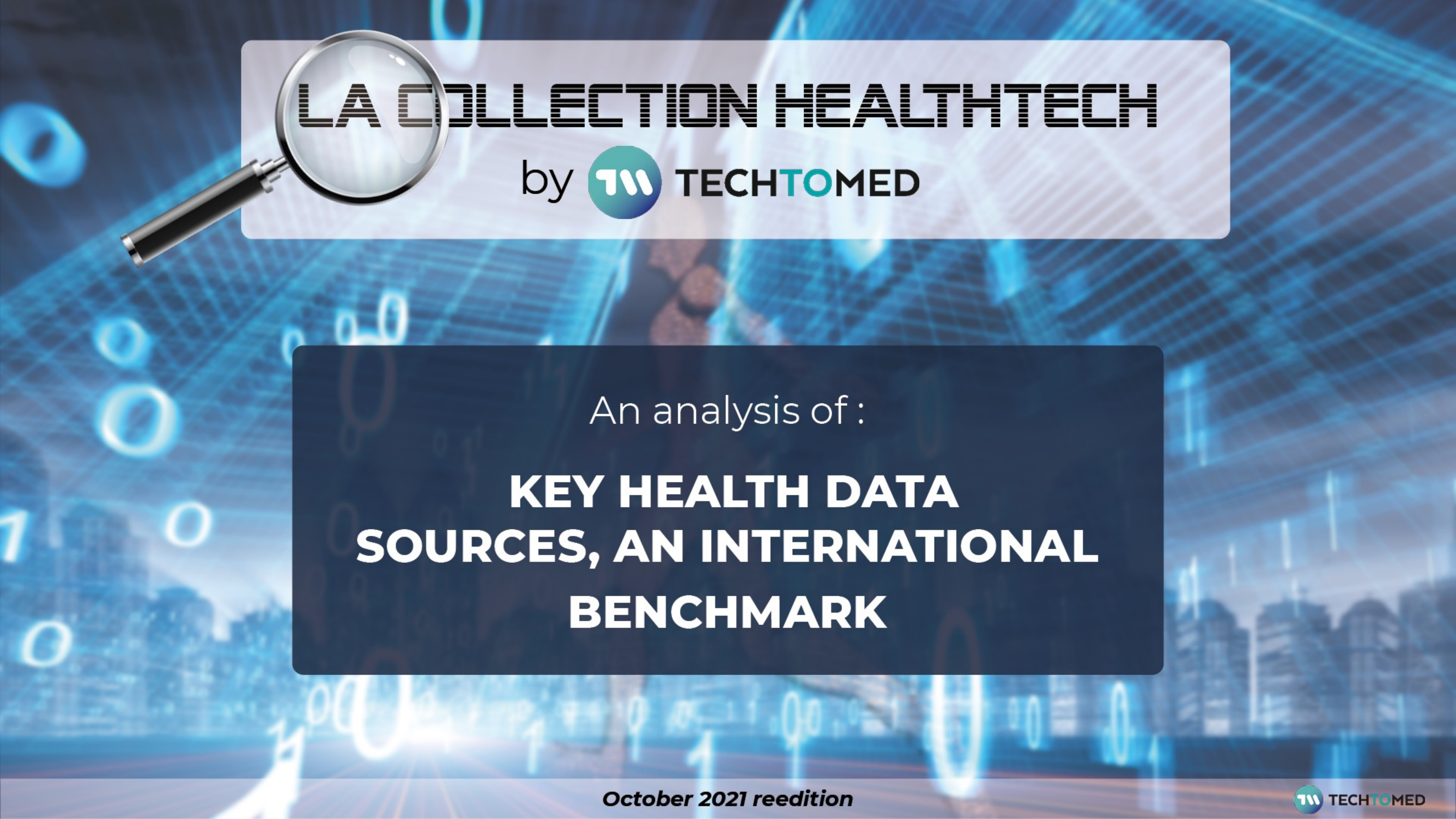
Percentage of population for which a particular drug in a class is ineffective



According to : Personalized Medicine Coalition

With personalized medicine, access to patient data allows healthcare providers to design individualized treatments that improve outcomes.





LA COLLECTION HEALTHTECH

by TECHTOMED

An analysis of :

KEY HEALTH DATA SOURCES, AN INTERNATIONAL BENCHMARK

October 2021 reedition