

Global Data Barometer

Health and Covid-19 module

Policy brief



Global Data Barometer



D4D.net
Data for Development



IDRC · CRDI

International Development Research Centre
Centre de recherches pour le développement international

Acknowledgements

Suggested citation: Global Data Barometer (2022). Health Module- Policy Brief – Global Data Barometer. ILDA

Author: Francesca De Chiara
Edition: Maria Esther Cervantes

Project direction: Silvana Fumega
Research and data coordination: Nicolas Grossman
Research and capacity building: David Zamora
Communication: Fiona Chawana with the key support of Cat Cortes and Violeta Belver

Executive summary

Data is critical for governments to properly manage crises like the COVID-19 pandemic. The adoption of exceptional measures to access and process datasets and establishing new patterns for publishing and communicating data, often including daily or weekly updates, is crucial in the fight against the diffusion of the virus. Health data have mainly been used on infection rates and mortality. Scientific data sharing is also vital to tackle the spread of the virus. The analysis conducted by the Global Data Barometer (GDB) considers three leading composite indicators focusing on vital statistics, real-time healthcare system capacity, and vaccines to answer a set of questions on the availability, accuracy, and timeliness of health and Covid-19 data. Public data availability, accuracy, and timeliness are indispensable elements in improving the response to the current emergency and future planning. The Covid-19 pandemic has dramatically led to public health systems' scrutiny by revealing gaps and disparities, unequal access to the healthcare system, uneven distribution of tests and vaccines, and improving marginalization.

Introduction

Public health work, research, and primary healthcare provision draw substantially upon data that should be reliable, accessible, and trusted. In the COVID-19 pandemic, governance and reusability of data in the health sector have been recognized as essential. The combination of respect for individual privacy rights and the focus on population and patient health is critical to ensure coordinated action. Data can support access to healthcare across the diverse systems of the world. Health data is essential for achieving goal 3 of the Sustainable Development Goals: "Ensure healthy lives and promote well-being for all ages." Achieving this goal requires understanding and improving population health.

Traditionally, population statistics are compiled through civil registration and vital statistics (CRVS) systems, censuses, and other surveys of individuals and households. In most countries, a population and housing census is conducted every ten years; in others, at irregular intervals of time¹. The census is also used to produce social statistics, but household surveys are usually a key data source for these statistics in many countries. However, the release of data and statistics from censuses and surveys often suffer from a substantial lag, affecting the need for timely evidence required by policymaking. Crises such as the COVID-19 pandemic, require real-time data, particularly on the population's location, density, and movements, while preserving personal data privacy and protection².

The civil registration and vital statistics (CRVS) systems have become increasingly recognized as catalytic for monitoring population dynamics spanning the entire life course. Population scientists have a long history of contributing to strengthening CRVS systems and using vital registration data³. The data provided by CRVS systems contains the information needed to understand population and development dynamics to achieve SDG 3

¹ Stats Brief April 2021, Issue no. 29: Big data for population and social statistics https://www.unescap.org/sites/default/d8files/knowledge-products/Stats_Brief_Issue29_Big_data_for_population_and_social_statistics_Apr2021.pdf

² United Nations ESCAP. Big data for population and social statistic. 2021. <https://repository.unescap.org/bitstream/handle/20.500.12870/3500/ESCAP-2021-PB-Big-data-population-social-statistics.pdf?sequence=1&isAllowed=y>

³ Silva, Romesh. "Population perspectives and demographic methods to strengthen CRVS systems: introduction." *Genus* 78, no. 1 (2022): 1-15. <https://link.springer.com/article/10.1186/s41118-022-00156-8>

by tracking the progress on mortality, providing data to improve vaccines, and strengthening the capacity to manage national and global health risks, ensuring an equal distribution of healthcare resources. Healthcare system capacity data supports governments and other actors — a crucial goal amid a worldwide public health crisis.

As noted in the World Health Organization (WHO) COVID-19 Strategic Preparedness and Response Plan⁴ (SPRP 2021), “healthcare systems and workers are under extreme pressure in many countries in terms of capacity and capabilities, financial resources, and access to vital commodities and supplies including medical oxygen”. In this context, real-time or very recent data about the healthcare system’s capacity improves and fastens decisions on how and where to direct patients and distribute resources more equitably. Data about healthcare system capacity and performance plays a crucial role in helping determine the quality of healthcare services. Based on the Global Data Barometer (GDB) survey results, we illustrate existing data gaps related to health and COVID-19, focusing on the three indicators related to the availability: vital statistics, real-time healthcare system capacity, and vaccination.

⁴ <https://www.who.int/publications/i/item/WHO-WHE-2021.02>

International Health and COVID-19 data gap analysis

According to the WHO guidelines, the states must collect and report aggregated data on COVID-19 confirmed cases and deaths, the number of tests, hospitalized cases, and discharges. On top of WHO data, the European Centre for Disease Prevention and Control (ECDC) asks European member states to collect and report data on the geographical spread of the pandemic; data related to COVID-19 healthcare impact (e.g., confirmed cases among healthcare workers, availability of ICU beds); outbreak surveillance data and contact tracing data. This aggregated data on COVID-19 hospitalized cases and discharges are often not publicly accessible or reported to ECDC or WHO. The data collection on COVID-19 confirmed cases or deaths among healthcare workers is not homogeneous. In most cases, the media reports outbreaks in nursing homes or workplaces, and contact tracing data are not publicly available.

The lack of stratification of COVID-19 cases and deaths by ethnic background, occupation, or socioeconomic status is also an issue, as stratification is needed to enrich previous minimum data sets (e.g., reporting cases and deaths by occupation, nationality, race, postal code) to uncover inequalities and achieve operational effectiveness in the epidemic's control at the local level. We also identify an urgency for investments in epidemiological surveillance and transparency. Without these, systems fail to collect adequate data and even report minimum indicators. These data gaps are mostly related to the lack of prioritization and under-investment in local public health surveillance systems. Given the time needed to achieve herd immunity through vaccination, immediate investment in comprehensive epidemiological surveillance is the only way to reduce transmission levels and avoid resurgences. We ask, how do governments make health and COVID-19 data available? It is evident from responses to the GDB survey that there is significant room for improvement. The GDB's research focuses on primary indicators and secondary data sources to assess the availability of vital statistics (primary indicator), healthcare system capacity data (secondary source), and real-time healthcare system capacity data (primary indicator). Regarding COVID-19, the data collection focused on indicators assessing the availability of vaccination data (primary indicator) and testing data (secondary source).

Vital statistics

The Vital statistics indicator focuses on the uses of CRVS systems for health. The research on the availability of vital statistics on a global scale has revealed the extent to which the civil registration and vital statistics (CRVS) information is available as open data. Among the uses of the CRSV Data, we highlight its foundational function for calculating and understanding excess deaths for large-scale diseases events such as COVID-19. What find that CRSV data should include the level of completeness by province, state, county, or other relevant regional category and list causes of death at least standardized to the 10th version of the International Classification of Causes of Death⁵ or another easily interoperable standard. Birth information should include details about the child, the birth, parents, and the registration process. Mortality data should comprise age, sex, geographic location, and cause of death.

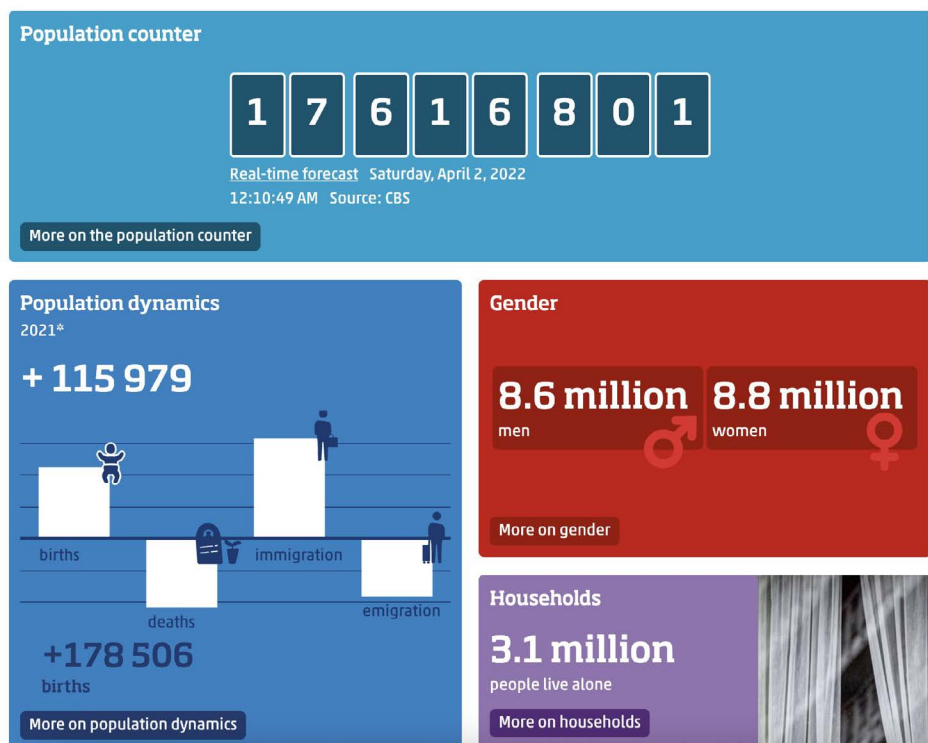
In 22% of the countries mandated data is unavailable online. Over 75% of the countries declare data is available from the government or government actions. For example, this data is closed in Estonia, where the government maintains the population register with accurate information on everyone living and residing within the country. Statistical data and aggregations based on this data can be found through the National Institute for health development (TAI) databases⁶. In the Netherlands, the Central Bureau of Statistics (CBS) provides data on the open data platform. CBS has been authorized to receive all data required to compile population statistics, including vital statistics. This aligns with requirements and standards from the United Nations Department of Economic and Social Affairs (UNDESA), Eurostat, and the Council of Europe. CBS has also developed a population data dashboard with basic population facts and figures. Via this dashboard, people access underlying datasets with specific data regarding mortality, gender, households, life expectancy, and migration.

⁵ WHO. International Statistical Classification of Diseases and Related Health Problems (ICD). <https://www.who.int/standards/classifications/classification-of-diseases>

⁶ Statistics Estonia. <https://www.tai.ee/et/statistika-ja-registrid> and Statistics Estonia - <https://www.stat.ee/>

Population dashboard

Up-to-date figures on the Dutch population.



Population dashboard - The Netherlands

Source: <https://www.cbs.nl/nl-nl/visualisaties/dashboard-bevolking>

Most countries provide national coverage. Findings show promising results but are not satisfying in the completeness of vital statistics. Regarding data fields and quality, 58% of the countries answered that the data includes no information on data limitations, specifically on the completeness of vital statistics in different provinces, counties, or regions. Birth data are often incomplete, lacking live-birth order and interval between last and previous live births to mothers (75% of the countries). As an example, Costa Rica⁷ also provides information on the father's age and place of usual residence. Other important data related to births such as place of occurrence, place of usual residence of mother, the month of the event, place of registration and month of registration, age, educational attainment, and ethnic and national group of the mother are not made available in more than half countries evaluated.

Data is available free of charge in 74% of the countries and openly licensed in 54% of cases. Also, less than half of countries provide data in machine-readable formats (JSON, CSV, xlsx, px, JSON-stat).

⁷ INEC Costa Rica. Nacimientos. <https://www.inec.cr/poblacion/nacimientos>

Real-time healthcare system capacity

Regarding the availability of real-time healthcare system capacity, the GDB's main findings reveal to what extent the information is available as open data. Real-time or near real-time data about the healthcare system's capacity should be available at the level of individual facilities and include details such as the number, type, and availability of beds, tests, vaccines, and devices. This level of detail is crucial for managing a pandemic in different ways. It can also benefit and impact decisions at different levels and different types of actors.

The need for this data is evident in cases of urgent need to identify locations to receive treatments and to get vaccines; determine inequalities in the healthcare system affecting members of marginalized communities; ease decisions of the government officials on where and when to build surge or overflow facilities; manage healthcare providers' access of patients with COVID-19 symptoms to facilities by knowing the current capacity to treat them. The availability of such data fastens the interventions, makes a wide range of actors understand the scarcity of resources, and facilitates decisions oriented to filling the gaps in personal protective equipment (PPE), medical devices, vaccines, etc. This indicator emerges from data practices observed with the COVID-19 pandemic. The GDB has investigated the findability of data on COVID-19-related websites run by the government, academia, or civil society.

There is a wide variety in the type of data published by each country. For example, Denmark, provides data on Covid-related patients⁸ and weekly reports on capacity. In India, Andhra Pradesh provides data on ICU beds⁹ and ventilators Germany also provides data on ICU beds¹⁰. Data on Covid Tests are available in Hong Kong and Taiwan¹¹. In France, you can navigate updated data on vaccinations¹². The Covid 19 data portal of New Zealand is maintained by

⁸ Covid-related patients: <https://www.sst.dk/da/corona/COVID-19-og-ny-coronavirus/Coronatal>; Weekly reports on (covid) capacity: <https://www.sst.dk/da/corona/Sundheds-og-aeldresektoren/Sygehusdrift>

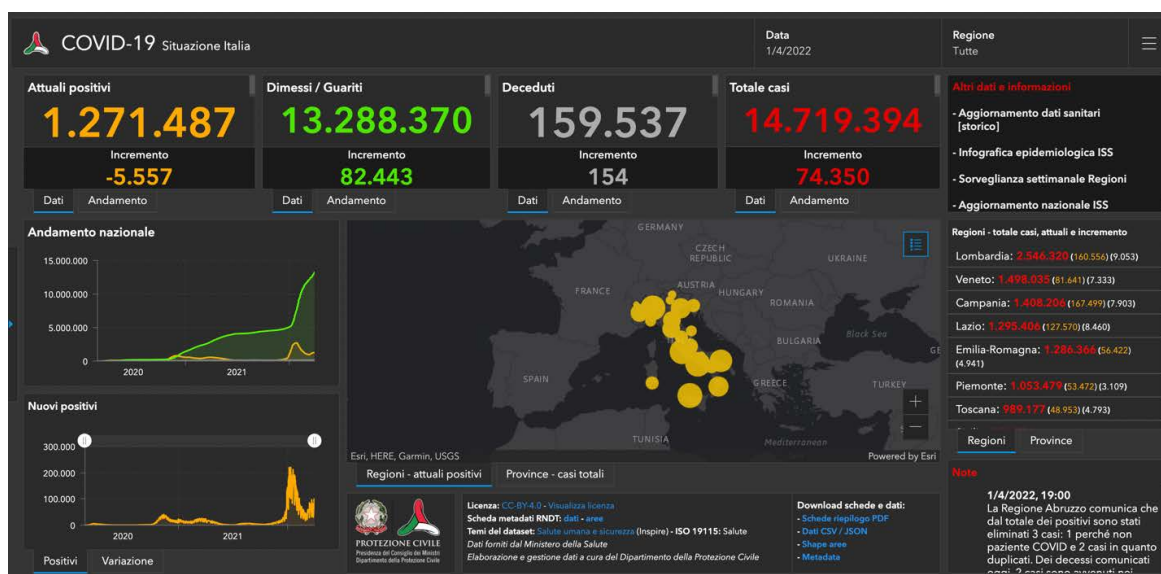
⁹ Andhra Pradesh. ICU beds http://dashboard.covid19.ap.gov.in/ims/hospbed_reports/

¹⁰ <https://www.intensivregister.de/#/aktuelle-lage/kartenansichten>

¹¹ Government of Hong Kong. COVID Tests <https://www.communitytest.gov.hk/en/> and Government of Taiwan <https://data.gov.tw/dataset/120450>

¹² Republic of France. Données relatives aux stocks des doses de vaccins contre la COVID-19 <https://www.data.gouv.fr/fr/datasets/donnees-relatives-aux-stocks-des-doses-de-vaccins-contre-la-covid-19/>

Stats NZ and reports on economic, environmental, and social aspects of COVID-19's impact on New Zealand and its recovery. Data visualizations on hospitals' utilization¹³ and capacity status in the United States are available on a dedicated website. The examples are the exception to the rule. The situation on a global scale is not encouraging, and there's much to do in terms of providing real-time data on the healthcare system's capacity.



Italian Covid-19 Dashboard. Source: Protezione Civile

<https://opendatadpc.maps.arcgis.com/apps/dashboards/b0c68bce2cce478eac82fe38d4138b1>

As an example, the Department of Civil Protection of Italy does daily updates of a dashboard under a CC-BY-4 license. The data they publish is: currently positive, total people now positive both hospitalized and in isolation; total people clinically healed; deceased persons (awaiting ISS verification); total people who tested positive. Total and current positive tests can be filtered by region and province. Situation maps by province show the total number of people who tested positive. Graphs show the national trend of the current positives, the healed and the deceased, and the daily increase. Data download links are available in Italy in various formats.

¹³HHS Protect Public Data Hub. Hospital Utilization <https://protect-public.hhs.gov/pages/hospital-utilization>

Vaccination (COVID-19)

Data about COVID-19 vaccines and vaccination should include the number and type of doses available and what percentage of total doses this represents. It should provide basic demographic information about vaccinated people, such as age, sex, disability, and membership in a marginalized population. It should also include information regarding the location and how the process is distributed. We should be able to use data about vaccines and vaccination to track the progress of vaccination and immunization; assess and update public health restrictions; identify disparities in the vaccination's occurrence; to support the analysis of the duration, efficacy, and other aspects of the various vaccines; to plan for the eradication of COVID-19.

An effective response to the health crisis necessitates timeliness and accuracy of vaccination data. However, there are several limits that affect them. A few countries that handle the production of vaccines have hoarded vaccines, causing unevenness in the global vaccine rollout. New variants of the virus and scarcity of information on the duration and efficacy of the existing vaccines are critically decreasing the quality of vaccination data. As with testing and treatment, access to vaccines has proven to be frequently inequitable. Disparities are likely to track the social determinants of health; and may be correlated with location or demographic variables. We have research highlights on some countries: the COVID-19 vaccination data by El Salvador¹⁴; the UK vaccination data¹⁵, including England, Scotland, and Wales with more detail available from the UK nation's health services and the Sri Lankan web page with updated data¹⁶.

Relevant issues identified concern data on the status of vaccinations on the disabled, marginalized groups, or vaccinated residents of long-term care facilities. The data does not provide clear disaggregation between different age groups. In Italy, a Github repository¹⁷ has been created and is regularly updated with information on vaccine doses (how many are received and then administered, including their geographic distribution) for people who have been vaccinated (by age and gender and number of doses, now including data on third doses), as well as data on people who caught Covid but then recovered.

¹⁴ Gobierno de El Salvador. Situación Nacional COVID 19. <https://covid19.gob.sv/>

¹⁵ United Kingdom. Vaccinations in the United Kingdom. <https://coronavirus.data.gov.uk/details/vaccinations>

¹⁶ Epidemiology Unit of the Ministry of Health of Sri Lanka. Coronavirus 2020-2022. https://www.epid.gov.lk/web/index.php?option=com_content&view=article&id=225&lang=en

¹⁷ Italy. Open Data Vaccini. <https://github.com/italia/covid19-opendata-vaccini>

Conclusion and recommendations

The insights deriving from the analysis of the GDB's survey results focused on governments' responsibilities such as issuing, collecting, distributing, and publishing health and Covid-19 data. It is clear how important identifying essential information for monitoring and managing the pandemic is. It is crucial to highlight the minimum required and mandatory information and the optional information. It is important to identify the minimum geographical granularity among the minimal requirements. The main recommendation is to decide which information is essential to report on a dataset and the data that must necessarily be collected. Another important aspect is that the information is stable, guaranteeing data quality and a standard definition of the data fields.

To summarize, our recommendations are the following:

- Provide the minimal geographic granularity, the time interval (i.e. daily), the quality, and the number of information (data fields).
- Use of controlled vocabularies and definition of the age range the level of aggregation (i.e. age groups) Other examples: the type of disease (ie: Influenza, COVID,) essential information (sex, age range, any diseases of the patient.) Nice to have data would be ethnicity of ob type. Consider implementing mandatory data and optional data, setting geographic granularity, are the crucial elements to guarantee and optimize data management.
- Examine the impact on privacy of data collection, considering the local, national and supranational regulations.
- Use standardized definitions of the data breakpoints that allow for comparisons. This must be organized at a supranational level to enable coordinated action, the channels and how states can deliver data, and to whom.
- Define quality and timing, and consistent data revisions to guarantee data quality.
- Further research can benefit from studying the significant comorbidities that affect patients.

