World Health
Organization

Glossary of health data, statistics and public health indicators

World Health Organization

Glossary of health data, statistics and public health indicators

Glossary of health data, statistics and public health indicators

ISBN 978–92–4-010548–5 (electronic version)
ISBN 978–92–4-010549–2 (print version)

© World Health Organization 2025

Some rights reserved. This work is available under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 IGO licence (CC BY-NC-SA 3.0 IGO; https://creativecommons.org/licenses/by-nc-sa/3.0/igo).

Under the terms of this licence, you may copy, redistribute and adapt the work for non-commercial purposes, provided the work is appropriately cited, as indicated below. In any use of this work, there should be no suggestion that WHO endorses any specific organization, products or services. The use of the WHO logo is not permitted. If you adapt the work, then you must license your work under the same or equivalent Creative Commons licence. If you create a translation of this work, you should add the following disclaimer along with the suggested citation: "This translation was not created by the World Health Organization (WHO). WHO is not responsible for the content or accuracy of this translation. The original English edition shall be the binding and authentic edition".

Any mediation relating to disputes arising under the licence shall be conducted in accordance with the mediation rules of the World Intellectual Property Organization (http://www.wipo.int/amc/en/mediation/rules/).

Suggested citation. Glossary of health data, statistics and public health indicators. Geneva: World Health Organization; 2025. Licence: CC BY-NC-SA 3.0 IGO.

Cataloguing-in-Publication (CIP) data. CIP data are available at https://iris.who.int/.

Sales, rights and licensing. To purchase WHO publications, see https://www.who.int/publications/book-orders. To submit requests for commercial use and queries on rights and licensing, see https://www.who.int/copyright.

Third-party materials. If you wish to reuse material from this work that is attributed to a third party, such as tables, figures or images, it is your responsibility to determine whether permission is needed for that reuse and to obtain permission from the copyright holder. The risk of claims resulting from infringement of any third-party-owned component in the work rests solely with the user.

General disclaimers. The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of WHO concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted and dashed lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by WHO in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by WHO to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall WHO be liable for damages arising from its use.

Editing and design by Inis Communication

Contents

Acknowledgements vii

Abbreviations viii

Background and methods ix

Glossary terms in alphabetical order 1

A

Accountability 1
Accuracy 1
Administrative area 1
Age groupings 1
Age-specific mortality rate 1

B

Birth registration data 1

C

Care cascade 1
Catchment area 1
Causes of death 2
Chi-square (χ2) test 2
Civil registration and vital statistics 2
Completeness of reporting 2
Compliance 2
Composite indicator 2
Confidentiality 2
Confounding 2
Construct validity 2
Content validity 3
Convergent validity 3
Core indicator 3
Correlation analysis 3
Count 3
Covariates 3
Credibility 3
Criterion validity 3

D

Data 3
Data access control 3
Data accessibility 3
Data anonymization 3
Data collection level 3
Data collection method 4
Data consent 4
Data custodian 4
Data element 4
Data governance 4
Data harmonization 4
Data information pyramid 4
Data inputs 4
Data integration 4
Data integrity 4
Data life cycle 4
Data privacy 5
Data provider 5
Data standards 5
Data type 5
Data quality 5
Data quality assessment 5
Data quality assurance 5
Data security 5
Data source 5
Data steward 5
Data stewardship 6
Data triangulation 6

Death registration data 6
Death registration desk 6
Denominator 6
Descriptive analysis 6
Digital health 6
Disability adjusted life years (DALYs) 6
Disaggregation 6
Disease surveillance system 6
Domain 6

E

Ecological analysis 6
Ecological fallacy 7
Effective coverage 7
Effect size 7
Estimation method 7
Evaluation 7
Event-based surveillance 7
External consistency of data 7
External responsiveness 7

F

Feasibility 7
Focal point 7

G

Gender 8
Geocoding 8
Geo-enablement 8
Geographic accessibility 8
Geographic information systems 8
Geospatial analysis 8
Geospatial data 8
Global Health Observatory 8
Global health estimates 8
Gold standard 8
Granularity 8

H

Health and demographic surveillance system 8
Health data 9
Health estimates 9
Health facility census 9
Health facility survey 9
Health impact assessment 9
Health indicator 9
Health inequality 9
Health inequity 9
Health information system 9
Health management information system 9
Health survey 10
Healthy life expectancy at birth 10
Heaping of data 10
Home-based record 10

I

Impact indicator 10
Imputation 10
Incidence rate 10
Indicator-based surveillance 10
Indicator classification 10
Indicator definition 10
Individual health record 11
Input indicator 11
Interaction 11
Internal consistency of data 11
Internal responsiveness 11
International Classification of Diseases (ICD) 11
International Health Regulations 11
Interoperability 11
Intervention coverage 11

J

Joint external evaluation 11

K

Kappa statistic (K) 12

L

Linkage 12
Logic model 12

M

Master facility list 12
Mean 12
Measure 12
Measurement 12
Measurement level 12
Measurement method 12
Median 12
Medical certification of cause of death 12
Metadata 12
Method of aggregate estimation 12
Microdata 12
Monitoring 13
Monitoring and evaluation (M&E) framework 13
Morbidity data 13
Mortality coder 13
Mortality data 13

N

National health strategic plan 13
National statistics office 13
Nationally representative 13
Negative predictive value 14
Notifiable conditions 14
Null hypothesis 14
Numerator 14

O

Odds ratio 14
Outcome-adjusted coverage 14
Outcome indicator 14
Output indicator 14

P

Percentage 14
Periodicity 14
Population attributable fraction 15
Population-based survey 15
Population census 15
Positive predictive value 15
Predictive validity 15
Preferred data sources 15
Prevalence rate 15
Primary data 15
Process indicator 15
Process of validation 15
Processed health data 15
Proportion 16
Proxy health indicator 16
Public health surveillance system 16
Publicly available 16
Punctuality 16
P value 16

Q

Quality-adjusted coverage 16

R

Rate 16
Ratio 16
Rationale 16
Raw health data 16
Record linkage 16
Registrar 17
Register 17
Registry 17
Registration form 17
Regression 17

Relevance 17
Reliability of data 17
Representativeness 17
Response rate 17
Responsiveness 17
Results framework 17
Routine data 18
Routine health information system 18

S

Sample 18
Sample size 18
Sampling error 18
Secondary data 18
Sensitivity analysis 18
Sex 18
Stakeholder 18
Standard operating procedures 19
Statistical model 19
Statistical significance 19
Statistics 19
Stratification 19
Subnational 19
Surveillance 19
Survey 19
Survey data 19
Sustainability 19
Sustainable Development Goals (SDGs) 19

T

Target population 19
Timeliness 20
Tracer indicator 20
Transparency 20
Trend 20

U

Uncertainty measure 20
Understandable/Simplicity 20
Unit of measure 20
Usefulness/Utility 20

V

Validity 20
Verbal autopsy 20
Vital event 21
Vital statistics 21

W

Weighting 21

Y

Years of life lost 21
Years lived with disability 21

Glossary terms by related thematic areas **22**

References **26**

Acknowledgements

The entire process from conception to publication has been led by Theresa Diaz (Unit Head, Epidemiology, Monitoring and Evaluation, Maternal, Newborn, Child and Adolescent Health (MCA), and Ageing, World Health Organization (WHO) and Co-chair, Indicator Governance Working Group). The first draft of the glossary was developed by Debra Jackson, London School of Hygiene & Tropical Medicine (LSHTM) and Yemi Okwaraji, LSTMH. The following people contributed to the elaboration of the glossary, reviewing the terms and definitions, as well as to the dissemination and analysis of an online survey to obtain feedback on the glossary terms: Philippe Boucher, Department of Data and Analytics (DNA,WHO); Craig Burgess (DNA, WHO), Tashi Chozom (DNA, WHO), Diana Estevez (DNA, WHO), Gerard Lopez (MCA, WHO).

Indicator Governance Working Group members (all are also members of the WHO Data Hub and Spoke Collaborative) reviewed the terms and provided feedback at various intervals:

- WHO Department of Data and Analytics/Department of Delivery for Impact (DDI): Philippe Boucher, Zoe Brillantes, Anh Chu, Diana Estevez, Steve MacFeely and Jean-Francois Saint Pierre.
- WHO Regional Office for the Americas (Pan American Health Organization): Adrienne Cox, Andrea Gerger and Oscar Martinez.
- WHO Regional Office for the Eastern Mediterranean: Eman Aly and Henry Doctor.
- WHO Regional Office for South-East Asia: Amani Siyam.
- WHO Science Division: Carl Leitner, Natschja Nash-Mendez and Ritika Rawlani.
- WHO Universal Health Coverage/Healthier Populations Division (HEP): Rifat Hossain and Richard Kumapley.
- WHO Universal Health Coverage /Life Course Division (UHL): Jenny Cresswell, Theresa Diaz, Julien Dupuy, Marta Gacic-Dobo, Jotheeswaran Thiyagarajan, Kathleen Strong and Chelsea Taylor.
- WHO Universal Health Coverage/Communicable and Noncommunicable Diseases Division (UCN): Junerlyn Farah Agua, Leanne Riley and Pamela Mbabazi.

Other WHO Hub and Spoke Collaborative members and WHO staff have also contributed feedback:

- DDI: Roxanne Moore, Nargis Rahimi, Amy Sy, Dan-George Vasilache, Wendy Venter.
- UHL: Mathieu Bonoil, Khassoum Diallo, Cristin Fergus, Regina Guthold, Elizabeth Katwan.
- HEP: Elisa Dominguez, Richard Johnston, Yuki Minato.
- UCN: Hazim Timmi.

All external experts submitted a declaration of interest to WHO disclosing potential conflicts of interest that might affect, or might reasonably be perceived to affect, their objectivity and independence in relation to the subject matter of the toolkit. WHO reviewed each of the declarations and concluded that none could give rise to a potential or reasonably perceived conflict of interest related to the subjects discussed at the meeting or covered by the guidance.

Abbreviations

DALY	disability adjusted life year
HIS	health information systems
ICD	International Classification of Diseases
M&E	monitoring and evaluation
NSO	national statistics office
SDG	Sustainable Development Goal
WHO	World Health Organization
YLL	years of life lost
YLD	years lived with disability

Background and methods

It is critical to have a standard set of terminology for health statistics, data and public health indicators. The World Health Organization (WHO) has produced multiple documents in various areas of health that include an accompanying glossary of terms relevant for a given subject area. Unfortunately, specific terms may be defined differently across these products. Additionally, for health indicators that are maintained within the WHO environment there are metadata terms that need to be completed, and it is necessary to have a clear definition of each of these terms to ensure standardization. To provide a standardized definition for public health data, statistics and health indicators, a stand-alone glossary of terms has been developed. It is intended for data managers, health analysts, monitoring and evaluation and surveillance specialists and other stakeholders who deal with health statistics, data and public health indicators across the data cycle: from collection to use.

This glossary was created by first developing a list of health indicator terms in need of definitions and then canvassing WHO publications, other United Nations publications, and then examining governmental public health institutions' definitions to define each term. The terms and definitions were subsequently reviewed by the Indicator Governance Working Group, part of the WHO Data Hub and Spoke Collaborative. This working group comprises volunteer members from the broader collaborative, including data focal points from each department at WHO headquarters and data and delivery focal points from each WHO regional office.

In addition to the regular progress updates provided to the Data Hub and Spoke Collaborative, the glossary, once reviewed and agreed upon by the Indicator Governance Working Group, was distributed via an online survey to the larger WHO Data Hub and Spoke Collaborative in February 2024. Based on the survey feedback, terms and definitions were retained if 75% of respondents agreed with them and removed if 50% or more suggested their removal. Additionally, new terms and definitions were incorporated based on the suggestions received. It was also discovered that several other glossaries were being created within WHO on health data terms and it was decided to consolidate these efforts into one glossary to avoid duplication. Additional health data and statistical terms were added during this consolidation process. A review of the forthcoming WHO data strategy and the existing WHO Personal data protection policy were carried out to ensure consistency of definitions. The scope of this glossary was refined to include the following terms:

- Public health data terms: Including those related to collection, disaggregation, analysis and management of public health data.
- Statistical terms: Incorporating the most common statistical methods used to analyse public health data and health indicators.
- Health indicator terms: Comprising terminology related to health indicators such as definitions, classification and validation.

These were again shared with members of the Data Hub and Spoke Collaborative via an online survey, and also to external specialists in data, measurement, monitoring and evaluation from WHO advisory groups and the multi-partner Health Data Collaborative: a platform that brings together over 1000 members from over 430 organizations to strength national health information systems (HIS) through effective collaboration, enabling countries to better monitor and attain health-related Sustainable Development Goals (SDGs). Relevant WHO advisory groups include Child Health Accountability Tracking (CHAT), Mother and Newborn Information for Tracking Outcomes and Results (MoNITOR), Global Action for Measurement of Adolescent health (GAMA), TAG4ageing (Technical Advisory Group for Ageing) and QoC Lifecourse Measurement (Life Course Quality of Care Measurement Group). Their input was essential in refining, shaping, and finalizing the list of terms and definitions in this glossary.

All comments received during the second online survey were reviewed. There were 85 respondents who provided comments on 74 terms, and recommended an additional 20 terms of which 10 were added because they represented important public health data, statistical, and/or health indicator terms. The recommended changes were considered by the Indicator Governance Working Group and changed based on consensus. The revised definitions and new terms were reviewed again by the WHO Data Hub and Spoke Collaborative members in a virtual meeting, and finalized based on consensus.

The terms are first listed in alphabetical order and then according to related thematic areas.

Glossary terms in alphabetical order

A

Accountability
Answerability or legal responsibility for identifying and removing obstacles and barriers to health services. This should include responding to findings from monitoring and evaluation *(1)*.

Accuracy
The degree to which a measurement, test or procedure correctly reflect the true value or condition *(2)*.

Administrative area
Delineated geographical areas within a particular territory created for the purpose of administration. For example: county, district, province, state, region, subnational or national *(3)*.

Age groupings
The combining of ages into groups for the purpose of data analysis. These groupings usually capture a time interval representing a developmental stage in the life course of a human. The actual aggregation used depends on the purpose of the analysis and the sample size *(4)*.

Age-specific mortality rate
A mortality rate of a particular age group. The numerator is the number of deaths in that age group; the denominator is the number of persons in that age group in the population *(5)*.

B

Birth registration data
Data collected in the process of recording a child's birth. It is a permanent and official record of a child's existence and provides legal recognition of that child's identity *(6)*.

C

Care cascade
A systematic framework used to track patient outcomes across the stages of care, from diagnosis to treatment and ultimately to successful outcomes. Each stage in the cascade identifies gaps in care that can prevent individuals from reaching the outcome such as receiving proper treatment or achieving disease control. The model was first developed in 1978 by Tanahasi to identify gaps and barriers in health systems performance *(7,8)*.

Catchment area
A geographical area delineated around an institution or business, such as a health facility, from where the population utilizes its services. Often a key output of geographic accessibility analysis (see also "geographic accessibility")*(9)*.

Causes of death

For the matters of international and national standard reporting, the definition is: The disease or injury that initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury. The fully specified title is: “Single underlying cause of death”. It is different from multiple cause analysis and other concepts like direct cause or cause determined by verbal autopsy. See also “International Classification of Diseases” *(10,11)*.

Chi-square (χ2) test

The Chi-square test of independence (also known as the Pearson Chi-square test) is one of the most useful statistics for testing hypotheses when the variables are nominal, as often happens in clinical research. Unlike most statistics, the Chi-square (χ^2) test can provide information not only on the significance of any observed differences, but also provides detailed information on exactly which categories account for any differences found *(12)*.

Civil registration and vital statistics

A system of continuous, permanent, compulsory, and universal recording of the occurrence and characteristics of vital events pertaining to the population, as provided through decree or regulation in accordance with the legal requirements in each country. The main source of vital statistics is civil registration and includes events like birth, marriage, divorce, adoption, death, and cause of death *(10)*.

Completeness of reporting

Reflects the percentage of reporting units that have provided data. This could be the completeness of facility reporting, district reporting or globally the completeness of reporting from countries *(13)*.

Compliance

The act of handling and managing personal and sensitive data in a way that adheres to regulatory requirements, industry standards and internal policies involving data security and privacy *(14)*.

Composite indicator

An indicator obtained by combining individual indicators from across several health topics based on an underlying model of the multi-dimensional concept being measured. Typically, a composite indicator measures a broader concept such as universal health coverage *(15)*.

Confidentiality

Refers to the ethical obligation of professionals, health institutions/ organizations and other health workers to preserve authorized restrictions on access to – and disclosures of – sensitive personal information about individuals *(16)*.

Confounding

Occurs when the relationship between an exposure and an outcome is distorted because the exposure is linked to other factors that also affect the likelihood of the outcome. A confounding variable is an additional variable that influences both the independent and dependent variables and that distorts the relationship between them *(17)*.

Construct validity

The degree to which a measurement aligns with related theoretical hypotheses, demonstrating that it accurately represents the underlying theoretical concept. It indicates how well the scores reflect the intended construct, ensuring the results are meaningful and relevant. Construct validity is commonly applied to process indicators *(18)*.

Content validity
The degree to which an assessment instrument is relevant to, and representative of, the targeted construct it is designed to measure *(19)*.

Convergent validity
How closely the indicator is related to other variables and other measures of the same construct. This approach is utilized when a gold standard does not exist. Convergent validity is often used for impact indicators *(20)*.

Core indicator
Core indicators may be defined in collaboration with all key stakeholders (e.g. ministry of health, national statistics office, other relevant ministries, professional organizations, experts, and major disease-focused programmes), and depends on the priority monitoring requirements related to health and health-related SDGs, among other health priorities *(21)*.

Correlation analysis
Examination of the strength and direction of linear relationships between two continuous variables *(22)*.

Count
A count gives the number of occurrences of the elements being measured (e.g. health event(s), facilities, health workers etc.), within a specified time and at a specified place. It is an integer, without any denominator *(23)*.

Covariates
Data, including non-health data, which are used in a statistical model to improve the estimation of the health indicator of interest. These data are population-specific and are available for every population included in the analysis. For example, a common covariate is gross domestic product per capita *(24)*.

Credibility
Confidence that users place in the statistics or data. Specifically, it refers to the believability or trustworthiness of the data *(2)*.

Criterion validity
The extent to which a measurement correlates with an external criterion of the phenomenon under study; ideally, a gold standard *(25)*.

D

Data
Basic, often unprocessed analogue or digital elements, characteristics or representations of facts or information. This includes raw, unprocessed numbers, measurements or text *(23,26,27)*.

Data access control
Access control is a security technique that regulates who or what can view or use data in a computing environment *(28)*.

Data accessibility
The ease with which users can find, retrieve, understand and use data *(29)*.

Data anonymization
The process of altering personally identifiable information within a dataset in such a way that the data subject is not or no longer identifiable *(30,31)*.

Data collection level
The specific setting from which the data are collected (e.g., national, subnational, facility, household, school or community) *(32)*.

Data collection method
Description of all methods used for data collection. This description should include, when applicable, the sample frame used, the questions used to collect the data, the type of interview, the dates/duration of fieldwork, the sample size, and the response rate *(33)*.

Data consent
Any freely given and informed indication of an agreement by the data subject to the processing of his/her personal data. It may be given either by a written or oral statement or other clear affirmative action *(30,31)*.

Data custodian
Agencies or individuals responsible for managing the use, disclosure and protection of source data used to create health indicators. Data custodians collect and hold information on behalf of a data provider or data steward *(34)*.

Data element
The smallest named item of data that conveys meaningful information *(35)*.

Data governance
The specification of decision rights and an accountability framework to ensure the appropriate behaviour in the valuation, creation, consumption and control of data and analytics. The rules, processes and behaviours related to the collection, management, analysis, use, sharing and disposal of personal and/or non-personal data *(28,36)*.

Data harmonization
All efforts to combine data from different sources and provide users with a comparable view of data from different studies or surveys. This is sometimes referred to as data integration *(37)*.

Data information pyramid
A schematic way of looking at the number of data items to be collected at each level of the health system allowing each level to gather data of importance and relevance to their daily work while avoiding excessive data collection where no action is taken. The pyramid illustrates how most data are collected at the base of the pyramid in the health facility, where most health service action takes place. Data are processed, filtered, and streamlined as data sets that are then passed up the health system *(38)*.

Data inputs
All numerical inputs to mathematical or statistical models that are used to generate global health estimates. Model inputs may include raw health data, processed health data, covariates and other parameters *(24)*.

Data integration
A process that merges multiple health data sources into a single, unified dataset. It involves combining data residing in different sources and providing users with a unified view of them *(28,39,40)*.

Data integrity
The extent to which all data are complete, consistent and accurate and are protected from deliberate bias or manipulation for political or personal reasons *(41,42)*.

Data life cycle
The main steps of the data life cycle include data collection, entry and recording, storage, processing and analysis, presentation and visualization, interpretation, sharing and dissemination, retention and archiving, maintenance and quality assurance, and disposal *(43,44)*.

Data privacy
The rights and practices that ensure individuals have control over how their personal information is collected, used, shared and maintained. It focuses on protecting personal data from misuse and ensuring that individuals can determine who has access to their data, how it is used, and for what purposes *(45)*.

Data provider
Individuals and organizations who are responsible, whether formally or informally, for making data accessible to others. Sometimes a data provider may be simply the producer of those data. In other cases, data may be deposited in a repository, centre or archive that has the responsibility of disseminating the data *(46)*.

Data standards
The rules for structuring information that has been collected, and which facilitate semantic interoperability. A set of agreed data standards ensures that the data entered into a system can be reliably read, sorted, indexed, retrieved and communicated between systems. For example, length of a field, format, the encoding of the information, and permissible values *(47)*.

Data type
Defines the form of a data element. These can include simple data types such as date, time, numeric value or complex data types, such as addresses *(48)*.

Data quality
Data quality is a set of standards that data should reach to be usable. Quality data must encompass the following characteristics: relevance, credibility, accuracy, timeliness, punctuality, methodological soundness, coherence and accessibility *(2)*.

Data quality assessment
The analysis or evaluation of data to determine its accuracy, completeness, consistency, and other quality attributes based on predefined criteria. This process helps identify issues and areas for improvement in the current state of data quality *(49)*.

Data quality assurance
The process of assessing and improving the quality of each data stream to ensure the overall quality of health data or health statistical products. Data quality assurance aims to ensure that the final data meets the required standards for accuracy, consistency and reliability *(49)*.

Data security
Appropriate organizational, physical and technical security measures implemented for both electronic and paper data to protect the security and integrity of personal data, including against personal data breach, and to ensure its continued availability for the purposes for which the personal data are processed *(31)*.

Data source
From where data are obtained (e.g. their origin such as the location setting and provider of the data). For example, routine health information systems, civil registration and vital statistics, population-based surveys, health facility surveys) *(50,51)*.

Data steward
Subject-matter experts who facilitate the use of data assets by all who need them, within the guardrails of relevant policy, and manage exceptions. Data assets do not belong to data stewards themselves, but they are responsible for taking care of them *(28,52)*.

Data stewardship
A collection of functions and practices that ensure all data and statistics of an organization are accessible, usable, safe and trusted *(52,53)*.

Data triangulation
Approach for critical synthesis of data from two or more sources to address relevant questions for programme planning and decision-making *(54,55)*.

Death registration data
Data collected during the civil registration of a death, including cause of death, sex, occupation, etc. *(11)*.

Death registration desk
A desk or office that receives death notification, validates information and enters it into a civil registration system *(11)*.

Denominator
The lower portion of a fraction; that is the number below the horizontal bar of a fraction. It is used to calculate, for example, rates, ratios and percentages. In a rate, the denominator is usually the population (or population experience, as in person-years, etc.) at risk *(5)*.

Descriptive analysis
A statistical method utilized to summarize and elucidate the primary features of a dataset. This approach allows for a clear understanding of the fundamental characteristics of the data without drawing inferences or making predictions *(56,57)*.

Digital health
The systematic application of information and communication technologies, computer science, and data to support informed decision-making by individuals, the health workforce and health systems, in order to strengthen resilience to disease and improve health and wellness *(58)*.

Disability adjusted life years (DALYs)
Represents the loss of the equivalent of one year of full health. For a disease or health condition it is the sum of the years of life lost to due to premature mortality (YLLs) and the years lived with a disability (YLDs) due to prevalent cases of the disease or health condition in a population *(59)*.

Disaggregation
Breaking down of observations to a more detailed level when finer details are required and made possible by the codes given to the primary observations. In health this often includes, sex, age, wealth quintile, education level, place of residence and occupation *(23)*.

Disease surveillance system
The infrastructure for ongoing systematic collection, analysis, storage and interpretation of outcome-specific disease data for use in planning, implementing and evaluating public health policies and practices *(60)*.

Domain
Categorization of health-related indicators into general groupings. For example, health status, risk factors, service coverage and health systems *(21)*.

E

Ecological analysis
Ecological analysis is a research methodology that investigates the relationships between different factors and health outcomes or other phenomena at the group level, rather than focusing on individuals. This approach entails analysing data gathered from populations or communities to uncover patterns, disparities, correlations, and possible causal relationships *(17)*.

Ecological fallacy
An erroneous inference that may occur because an association observed between variables on an aggregate level does not necessarily represent or reflect the association that exists at an individual level; a causal relationship that exists on a group level or among groups may not exist among the group individuals *(17)*.

Effective coverage
The proportion of individuals experiencing optimal health gains from a service among those who need the service. It is the fraction of potential health gain that is actually delivered to the population through the health system, given its capacity *(8,61,62)*.

Effect size
Refers to the magnitude of the expected difference between two groups in a study, which is expressed as an estimate. It is a crucial factor in determining the sample size needed for a study, with a smaller effect size requiring a larger sample size for detection *(63)*.

Estimation method
The steps taken to generate health estimates. An explanation of the method should include how the indicator is calculated, including mathematical formulas and descriptive information of computations made on the source data to produce the estimate (including adjustments and weighting). Reporting on methods used should align with the *Guidelines for accurate and transparent health estimates reporting* (GATHER) *(24)*.

Evaluation
A process that attempts to determine as systematically and objectively as possible the relevance, effectiveness and impact of activities in light of their objectives *(21)*.

Event-based surveillance
The organized and rapid capture of information about incidents that are a potential risk to public health. This information can be rumours and other ad hoc reports transmitted through formal channels (i.e. established routine reporting systems) and informal channels (i.e. media, health workers and nongovernmental organizations' reports) *(64)*.

External consistency of data
An assessment of the level of agreement between two sources of data measuring the same health indicator. The two sources of data that are usually compared are data flowing through health management information systems or a programme-specific information system, and data from a periodic population-based survey *(13)*.

External responsiveness
Reflects the extent to which change in a measure relates to corresponding change in a reference measure of clinical or health status *(65)*.

F

Feasibility
Refers to the availability of data to measure the indicator. If the data is available from existing health data at a reasonable cost and/or will not add additional burden to collect data *(66)*.

Focal point
The designated person to respond on behalf of a Member State or technical group during data collection or country consultation process *(67)*.

G

Gender
Refers to the characteristics of women, men, girls and boys that are socially constructed. This includes norms, behaviours and roles associated with being a woman, man, girl or boy, as well as relationships with each other. As a social construct, gender varies from society to society and can change over time *(68)*.

Geocoding
The process of transforming a description of a location – such as a pair of coordinates, an address, or the name of a place-to a location on the earth's surface *(69)*.

Geo-enablement
To apply geospatial capabilities to a business process in order to establish the authoritative spatial location of business data and enable contextual spatial analysis *(70)*.

Geographic accessibility
Geographic area and population living within it covered by a facility (e.g. health facility, school, etc.), calculated using mode of transportation *(71)*.

Geographic information systems
A computer system that creates, manages, analyzes, and maps all types of spatial data *(72)*.

Geospatial analysis
Geospatial analysis is the use of geographic information systems and geospatial data to measure, examine and visualize spatial patterns, relationships and trends. Often includes statistical methods and/or artificial intelligence *(73)*.

Geospatial data
Information describing the location and attributes of objects, events or other features on the earth's surface *(70,74)*.

Global Health Observatory
The WHO global database for health and health-related indicators, presenting the latest available data at global, regional and national levels. This database is the WHO gateway to health-related statistics, providing access to over 1000 indicators on priority health topics including mortality and burden of diseases *(75)*.

Global health estimates
WHO estimates on death and disability globally, by region and country, available by age, sex and cause. These provide key insights on mortality and morbidity trends and are a powerful tool to support informed decision-making on health policy and resource allocation *(76)*.

Gold standard
A method, procedure or measurement that is widely accepted as being the best available to test for or treat a disease or condition *(77)*.

Granularity
Granularity is the level of detail of the data *(78)*.

H

Health and demographic surveillance system
Longitudinal data collection platforms that track births, deaths, migrations and socioeconomic and health circumstances over time in established geographic areas *(79)*.

Health data
Raw, unprocessed numbers, measurements, that relate directly to the health and well-being status of an individual or to the health services that the individual receives *(23,80)*.

Health estimates
Quantitative population-level estimates (including global, regional, national or subnational estimates) of health indicators, including indicators of health status such as estimates of total and cause-specific mortality, incidence and prevalence of diseases, injuries and disability and functioning; and indicators of health determinants, including health behaviours and health exposures *(24)*.

Health facility census
Periodic enumeration of all public and private healthcare facilities within a country about the facilities and the services they provide *(81)*.

Health facility survey
Periodic enumeration of a representative sample of public and private health care facilities within a country, providing information about the facilities and the services they provide *(81)*.

Health impact assessment
A combination of procedures, methods and tools by which a policy, programme, product or service may be judged concerning its effects on the health of the population and the distribution of those effects within the population *(82)*.

Health indicator
A measurable quantity that can be used to describe a population's health or its determinants *(23,24)*.

Health inequality
A measured difference in health between population subgroups. Health inequalities can be measured and monitored *(83)*.

Health inequity
Unfair, avoidable or remediable differences in health among groups of people. In some cases, the absence of a difference between groups (that is, a situation of equality) might be considered inequitable. Health inequity is rooted in the unfair distribution of, and access to, power, wealth and other social resources, and is linked to forms of disadvantage that are socially produced, such as poverty, discrimination and lack of access to services or goods *(83)*.

Health information system
A system that collects data from health and other relevant sectors, analyses the data and ensures their overall quality, relevance and timeliness, and converts the data into information for health-related decision-making. It has four key functions: (i) data generation; (ii) compilation; (iii) analysis and synthesis; and (iv) communication and use. A solid health information system will be capable of generating reliable data from hospitals, outpatients, reportable diseases registries, surveys, civil registration and vital statistics and other relevant data for health *(23,84)*.

Health management information system
An information system specially designed to support planning, management and decision-making in health facilities and organizations *(85,86)*.

Health survey

A survey that is designed to gather information about health (physical and mental) and health-related factors. Health surveys generally include measures of risk factors, health behaviours, and non-health determinants or correlates of health such as socioeconomic status. The range of measures that can be included is wide and varies by survey. Age, sex/gender and race/ethnicity are the basic demographic variables that are included in health surveys. Socioeconomic determinants of health include education, income, geographic region and rural/urban residence *(87)*.

Healthy life expectancy at birth

Average number of years that a person can expect to live in 'full health' by taking into account years lived in less than full health due to disease and/or injury *(88)*.

Heaping of data

A measure of the tendency of respondents to report numerical values using preferred numbers or specific intervals when providing estimates, such as age of disease diagnosis or date of events (e.g. birthweights reported on 2000g or 2500g) or rounded (i.e. ending in '00' or '50') *(89)*.

Home-based record

A health document used to record the history of health services received by an individual. It is kept in the household, in either paper or electronic format, by the individual or their caregiver and is intended to be integrated into the health information system and complement records maintained by health facilities *(90).*

I

Impact indicator

Measures the ultimate objective that programmes are designed to affect, such as decreases in mortality and morbidity. Sometimes referred to as long-term outcome *(21,50)*.

Imputation

A method for retaining the majority of the dataset's data and information by substituting missing data with an inferred value *(91)*.

Incidence rate

A new event or case of a disease (or a death or other health condition) that occurred in a specified time period divided by the population at risk for the event in that time period *(23)*.

Indicator-based surveillance

Routine reporting of cases of a defined list of diseases, including notifiable disease surveillance systems, sentinel surveillance, and laboratory-based surveillance. Indicator-based surveillance commonly comes from health care facilities and can be regularly reported *(64)*.

Indicator classification

The level of measurement provided by the indicator; this can be one of five levels starting with input (lowest) and moving through process, output, outcome and finally impact (highest) *(21)*.

Indicator definition

How the indicator is measured, including numerators, denominators, data type and disaggregation in common use. The indicator definition should be unambiguous and be expressed in universally applicable terms *(23,92)*.

Individual health record
Files or a document that contain diagnoses and treatment, medications, allergies, immunizations, as well as radiology images and laboratory results for an individual. These records contribute to tracking a patient's medical history. They are generally kept in a health facility, in either paper or electronic format *(93)*.

Input indicator
Measures human and financial resources, physical facilities, equipment and operational policies that enable programme activities to be implemented. This includes health financing, health workforce, health infrastructure, and health information and governance *(21,50)*.

Interaction
When the relationship between two variables depends on the value of another variable *(17)*.

Internal consistency of data
A measure based on the correlations between different items on the same test or survey. It measures whether several items that propose to measure the same general construct produce similar scores or results. For example, if a respondent answered "Do you consider yourself healthy?" as yes, and then in the same survey answered "Do you consider yourself unhealthy?" as no, that is good internal consistency *(94)*.

Internal responsiveness
Characterises the ability of a measure to change over a prespecified time frame *(65)*.

International Classification of Diseases (ICD)
A structured translation of each medical condition into an alphanumeric code, which allows for the harmonization and comparison of mortality statistics across time and locations. ICD is a gold standard coding system for reporting cause of death data *(10)*.

International Health Regulations
It is an instrument of international law that is legally-binding on 196 countries, including the 194 WHO Member States. The International Health Regulations grew out of the response to deadly epidemics that once overran Europe. They create rights and obligations for countries, including the requirement to report public health events *(95)*.

Interoperability
The ability of different applications to access, exchange, integrate and use data in a coordinated manner through the use of shared application interfaces and standards, within and across organizational, regional and national boundaries, to provide timely and seamless portability of information and optimize health outcomes *(58)*.

Intervention coverage
People receiving a specific service (i.e. treatment, care, activity to improve health) among those who need it. It requires a well-defined service that can be measured, and precise measurement of the population in need of that service *(96)*.

J

Joint external evaluation
Voluntary, collaborative, multisectoral process to assess country capacities to prevent, detect and rapidly respond to public health risks *(97)*.

K

Kappa statistic (K)
A statistical measure of inter-rater reliability for categorical variables *(98)*.

L

Linkage
The process of combining data from different sources that relate to the same person to create a new, enhanced data resource *(99)*.

Logic model
A diagram that identifies and illustrates the linear relationships from programme inputs, processes, outputs, and outcomes to the impact *(100,101)*.

M

Master facility list
The unique, complete, up-to-date and uniquely coded list of all the active and prior health facilities in the country that are officially curated by the mandated agency. At a minimum, the list includes a unique ID, location, type, and name of each facility *(103)*.

Mean
The average of a set of values *(103)*.

Measure
A standard unit used to express the size, amount, or degree of something *(23)*.

Measurement
Refers to the extent, dimension, quantity, etc. of an attribute *(23)*.

Measurement level
The specific setting that the indicator is designed to measure/monitor (e.g. global, national, subnational, facility, household, community, school) *(21,23,50)*.

Measurement method
How the data from data sources are used; this can be process or other types of analyses that make use of the indicator *(104)*.

Median
Middle point of a set of ordered numbers; half of the values are higher than the median, and half of the values are lower *(17)*.

Medical certification of cause of death
Describes the process of accurately identifying the sequence of events that lead to death. It includes also recording of causes that contributed to death but were not part of the sequence that led to death *(10)*.

Metadata
Data that define or describe other data. They are the information needed to explain and understand the indicator or values being presented *(105)*.

Method of aggregate estimation
Description of the methodology, including any mathematical formulas, used for the calculation of the regional/global aggregates from country values *(106)*.

Microdata
Microdata are unit-level data obtained from sample surveys, censuses and administrative systems. They provide information about characteristics of individual people or entities such as households, business enterprises, facilities, farms or even geographical areas such as villages or towns *(107)*.

Monitoring
The systematic and routine collection of information to assess performance and progress towards specific targets and over an established period of time *(108)*.

Monitoring and evaluation (M&E) framework
A structured and systematic tool used in project management and programme implementation to assess performance, measure outcomes, and ensure the achievement of objectives. It consists of several core components, including clear project objectives, key performance indicators, data collection methods, data sources, and responsibilities. It defines how data will be collected, analysed and reported, ensuring that the project or programme remains on track *(100,101)*.

Morbidity data
The information registered on the state of being symptomatic or unhealthy due to a disease or health-related condition *(110)*.

Mortality coder
The trained person that registers medical conditions and events reported in Medical Certificate of Cause of Death forms to determine the underlying cause of death and assign mortality codes using ICD rules and principles *(10)*.

Mortality data
Information on the number of deaths caused by an event or illness in a specified time period. Measuring how many people die each year and why they have died is one of the most informative ways of assessing the effectiveness of a country's health system. These data allow health authorities to evaluate how they prioritize public health programmes *(10)*.

N

National health strategic plan
A plan that includes a set of priorities and activities to achieve key milestones that will have impact beyond the health sector. It includes medium- and long-term expected outcomes, and concrete and realistic allocation of resources to implement the activities within clear timing. The plan concretizes priorities; keeps focus on medium- and long-term goals without deviating from an optimal path. It integrates the health sector; helps focus the policy dialogue on health priorities; and guides operational planning, resource allocation and health sector monitoring and evaluation *(110)*.

National statistics office
Government agency or institution responsible for collecting, analysing, using and disseminating statistical data related to a country *(111)*.

Nationally representative
A survey or surveillance system that will use design methods and standardized criteria scalable to national context using a sub-sample that represents the target population in terms of age, sex, urban/rural and other categories of interest *(112)*.

Negative predictive value

The probability that following a negative test result a given individual will truly not have that specific disease. It is defined as the number of true negatives (people who test negative who don't have a condition) divided by the total number of people who test negative *(113)*.

Notifiable conditions

A disease that, when diagnosed, requires health providers (usually by law) to report to national, state or local public health officials. Notifiable diseases are of public interest due to reasons including their contagiousness, severity or frequency *(114)*.

Null hypothesis

The claim that the effect being studied does not exist. It can also be described as the hypothesis in which no relationship exists between two sets of data or variables being analysed *(115)*.

Numerator

Count of values captured by the indicator in a specified population. The upper portion of a fraction used to calculate a rate or ratio *(5)*.

O

Odds ratio

A measure of association between an exposure and an outcome. It represents the odds (or probability) that an outcome will occur given a particular exposure, compared to the odds of the outcome occurring in the absence of that exposure. The numerator is the proportion of the event of interest, and the denominator is the proportion of the non-event *(23)*.

Outcome-adjusted coverage

The proportion of a population that benefits from an intervention after considering the quality or effectiveness of the intervention. In this measure, coverage is not just about access to services but also reflects whether the intervention had a positive health impact. It adjusts traditional intervention coverage estimates by incorporating the actual outcomes of the health services provided *(8)*.

Outcome indicator

Measures whether the programme is achieving the expected effects/changes in the short, intermediate and long term, such as changes in intervention coverage or health-related behaviours. Some programmes refer to their longest-term/most distal outcome indicators as impact indicators *(21)*.

Output indicator

Measures the immediate products provided or services delivered as a result of the processes conducted in a programme or project, such as number of persons trained, number of children vaccinated *(21,116)*.

P

Percentage

Number or ratio that can be expressed as a fraction of 100 *(117)*.

Periodicity

Data can be compiled continuously in systems such as civil registries, cancer registries and surveillance systems of reportable diseases. Data can also be compiled periodically, which is to say at regular intervals or without predefined periodicity, and at a particular point in time *(23)*.

Population attributable fraction
The proportional reduction in population disease or mortality would occur if exposure to a risk factor were reduced to an alternative ideal exposure scenario. The number of deaths and DALYs attributed to a risk factor is quantified by applying the population attributable fraction to the total number of deaths or the total burden of disease *(118)*.

Population-based survey
Descriptive cross-sectional epidemiological study that is useful for calculating the prevalence of self-reported events or events measured during the investigation, generally employing a representative sample from the population of interest *(23)*.

Population census
Is the total process of planning, collecting, compiling, evaluating, disseminating and analysing demographic, economic and social data at the smallest geographic level pertaining, at a specified time, to all persons in a country or in a well-delimited part of a country *(119)*.

Positive predictive value
Is the probability that following a positive test result, a given individual will truly have that specific disease. It is defined as the number of true positives (people who test positive and have a condition) divided by the total number of people who test positive *(113)*.

Predictive validity
The degree to which predictions are confirmed by facts expressed in terms of its ability to predict future outcomes or events. Predictive validity is often used for impact indicators *(25)*.

Preferred data sources
The recommended source of data for a specific health indicator. This can include: civil registration and vital statistics systems; surveillance; national population-based surveys; routine facility information systems; health facility assessments; administrative data sources; human resources information systems; laboratory information systems *(50,51)*.

Prevalence rate
Is the proportion of people in a population who have a particular disease or attribute at a point in time or over a specified period of time *(120)*.

Primary data
Data that are gathered first-hand for a specific purpose. Data are reported as is, or with modest adjustment *(121)*.

Process indicator
Measures a programme's activities. This indicates whether the programme is being implemented as planned (e.g. health workforce training, constructing a health facility, the process of registering births and deaths)*(21,116)*.

Process of validation
A description of how the indicator was validated to assess how accurately it measures what it is intended to do. This glossary includes some methods for validation *(122)*.

Processed health data
Health statistics that have been calculated from raw health data, but are not the result of synthesizing multiple data sources. Examples of processing raw health data include cleaning data by removing implausible values, calculating an indicator with an algorithm, or adjusting a statistic for bias *(24)*.

Proportion
The size, number or amount of one thing or group as compared to the size, number or amount of another, when the numerator is a subset of the denominator *(23)*.

Proxy health indicator
An indicator that stands in for another indicator or topic that is difficult to measure or for which data are limited *(83)*.

Public health surveillance system
Is the system that systematically collects, analyses, and interprets health-related data essential to planning, implementation and evaluation of public health practice *(123)*.

Publicly available
Information in any form that is generally accessible, without restriction, to the public *(124)*.

Punctuality
The time lag between the release date of data and the target date on which they were scheduled for release as announced in an official release calendar *(2)*.

P value
The probability that a statistic would be as extreme as or more extreme than observed if the null hypothesis were true *(17)*.

Q

Quality-adjusted coverage
An assessment metric that evaluates how effectively a health intervention reaches a target population and the quality of care that is provided. This measure combines intervention coverage rates with quality measures to better understand the overall impact of health programmes *(8,125,126)*.

R

Rate
An expression of the frequency with which an event occurs in a defined population, usually in a specified period of time. The components of a rate are the numerator, the denominator, the specified time in which events occur, and usually a multiplier, that converts the rate from a fraction or decimal to a whole number. The numerator is the absolute number of occurrences of the event being studied in a specified time. The denominator is the reference population (or population being studied) at the same time *(17,23)*.

Ratio
The result of dividing one quantity by another without regard for details such as a time dimension. Rates, proportions and percentages are all types of ratios. The distinction between a proportion and a ratio is that, whereas the numerator of a proportion is included in the population defined by the denominator, this is not necessarily so for a ratio, which expresses the relationship of two separate and distinct quantities, neither of which is included in the other *(17)*.

Rationale
Importance of the indicator for public health response *(127)*.

Raw health data
Measurements derived from primary data collection with no adjustments or corrections *(23,80)*.

Record linkage
The methodology of bringing together corresponding records from two or more files or finding duplicates within files *(128)*.

Registrar
The official authorized by law to register the occurrence of vital events and to represent the legal authority of government in the field of civil registration *(11,129)*.

Register
A list or file containing uniform information about individual persons, collected in a systematic and comprehensive way, in order to serve a predetermined purpose *(130)*.

Registry
A collection of information about individuals, usually focused on a specific diagnosis or conditions. Registry data is stored in a database and can provide health care providers and researchers with first-hand information about people with specific conditions, both individually and as a group, and over time. Patient registries are different from health facility registers in that the former collect patient information that is disease- or condition-specific whereas the later collect information about the patient regardless of their condition or disease *(130)*.

Registration form
The registration form is the paper or electronic format that is used to register a vital event *(131)*.

Regression
A statistical technique that relates a dependent variable to one or more independent (explanatory) variables *(17)*.

Relevance
The degree to which the data meet the user needs. Indicators must provide information that is appropriate and useful for guiding policies and programmes as well as for decision-making *(2,23)*.

Reliability of data
The degree to which the results obtained by a measurement/procedure can be replicated. Consistency of the data when collected repeatedly using the same procedures and under the same circumstances [synonym: replicability] *(13,23)*.

Representativeness
The ability for the indicator to accurately describe the occurrence of a health-related event over time and its distribution in the population by place and person. This involves the absence of selection bias with respect to the population that the indicator is intended to represent *(23,66)*.

Response rate
The number of people who answered the survey divided by the number of people in the sample *(132)*.

Responsiveness
An indicator's ability to detect changes over time in response to interventions, treatments, or natural progression of the condition (see also: "External responsiveness" and "Internal responsiveness" *(133)*.

Results framework
A diagram that maps the direct causal relationships between incremental steps or results of key activities all the way up to the overall objective and goal of the programme or intervention. It includes an overall goal, strategic objectives, and intermediate results *(100,101)*.

Routine data
Data generated as part of regular operational or administrative processes within the health care system. These data are systematically collected during the delivery of health services and the management of health facilities. Routine data includes patient demographics, health conditions, treatments and outcomes recorded in health facility registers, electronic health records, and health management information systems. It also covers information related to health service resources, such as infrastructure, human resources, and financial transactions. The primary purpose of routine data is to support daily operations, monitor service delivery, and facilitate decision-making across various levels of the health system. This is also referred to as administrative data *(134,135)*.

Routine health information system
A system that generates data collected at public and private health facilities and institutions, and at community-level health care posts and clinics – at regular intervals of a year at minimum. The data give a picture of health status, health services and health resources. Most of the data are gathered by health care providers, by supervisors and through routine health facility surveys. The sources of those data are generally individual health records, records of services delivered, and resource health records including financial, commodity or laboratory records *(51,86)*.

S

Sample
Sample is a subset of a population *(136)*.

Sample size
Refers to the number of observations or data points gathered from a population for statistical analysis. It is a critical element of research design, significantly affecting the reliability and validity of a study's findings *(17)*.

Sampling error
That part of the difference between a population value and an estimate thereof, derived from a random sample, which is due to the fact that only a subset of the population is enumerated *(137)*.

Secondary data
Data that was originally collected for other purposes. The data from these existing sources are considered secondary. Although these data were not created for the purpose at hand, they facilitate the development of the required indicators. Data from a census, research, information system, etc. are secondary data *(23)*.

Sensitivity analysis
A systematic approach to evaluate how the variation in the output of a system or model can be attributed to different sources of variation in its inputs. It involves examining the sensitivity of the model's outcomes or outputs to changes in individual input parameters, providing insights into the relative importance of each parameter in influencing the overall results *(138)*.

Sex
The biological and physiological characteristics of females, males and intersex persons, such as chromosomes, hormones and reproductive organs *(68)*.

Stakeholder
Interested parties, group or organization who may affect, be affected by, or perceive itself to be affected by a decision, activity or outcome of a project or programme *(139)*.

Standard operating procedures
A set of step-by-step documented instructions created by an organization to ensure consistency in carrying out tasks and that standards are followed *(140)*.

Statistical model
A mathematical model that embodies a set of statistical assumptions concerning the generation of sample data (and similar data from a larger population) *(141)*.

Statistical significance
Mathematical measure of the probability that a result is likely due to chance or another factor. That is the probability of rejecting the null hypothesis when the null hypothesis is actually true *(142)*.

Statistics
The science of using information discovered from collecting, organizing and studying numbers *(26,27,143)*.

Stratification
The process of sorting data of subjects into defined segments or groups. This method can be used when sampling a population for conducting a survey or this can be used for analysis to control for confounding *(23)*.

Subnational
Geographical unit below national level. Subnational levels may indicate administrative divisions such as region, district or municipality levels, but may also indicate other divisions such as health districts.

Surveillance
The continuous, systematic collection, analysis, interpretation, and dissemination of data needed for the planning, implementation and evaluation of public health actions. Some examples are public health surveillance system, indicator-based surveillance, disease surveillance system, demographic surveillance system, etc. *(144)*.

Survey
A structured list of questions that collect data on a specific population *(117)*.

Survey data
A survey is an investigation about the characteristics of a given population by means of collecting data from a sample of that population and estimating their characteristics through the systematic use of statistical methodology *(145)*.

Sustainability
The source's potential to remain relevant and be of the quality needed to generate information over time. This depends not only on the periodicity of the data collection, but on the availability of the financial resources needed to sustain that source of data, the presence of a legal framework, political will, among other factors *(23)*.

Sustainable Development Goals (SDGs)
These are 17 global objectives that were agreed by Member States of the United Nations and aim to transform our world. They are a call to action to end poverty and inequality, protect the planet, and ensure that all people enjoy health, justice and prosperity. It is critical that no one is left behind *(104)*.

T

Target population
A population about which information is wanted and estimates are required. It can be a group of individuals who meet specific criteria for inclusion in a survey. In relation to health care, it is often a group of individuals in need of a specific intervention or service *(146)*.

Timeliness
When the data is quickly available and accessible for use. In the context of data quality, the degree to which reports are submitted on time according to established deadlines. Timeliness involves the availability and reliability of the data at the time it is needed to construct related indicators. Thus, timely indicators provide better opportunities for making health-related decisions *(13,23,66)*.

Tracer indicator
A highly specified indicator chosen as an example to represent a broader health topic *(83)*.

Transparency
Any information or communication relating to the processing of personal data is easily accessible and easy to understand, and that clear and plain language be used. Individuals should be made aware of risks, rules, safeguards and rights in relation to the processing of personal data and how to exercise their rights in relation to such processing *(147)*.

Trend
A pattern found in multiple standardized measurements over time; it is used to describe if the data is showing an upward or downward movement for a specified time period *(148)*.

U

Uncertainty measure
Relates to how the estimate might differ from the “true value”. Common measures of uncertainty include standard error, standard deviation, confidence interval, coefficient of variation and statistical significance *(149)*.

Understandable/Simplicity
When an indicator is presented in a clear, concise and easily comprehensible way. The indicator must be understood by those responsible for taking action and, specifically, by those responsible for decision-making *(23,60)*.

Unit of measure
Unit of measure in which the indicator is presented (e.g. deaths per 1000 live births; US$; litres per person per year).

Note: Percentage is not considered a unit of measure and indicators that are presented as percentages should have the unit of measure field filled in as ‘n/a’ (not applicable)

Usefulness/Utility
If the indicator is useful for programme improvement and policy issues *(60)*.

V

Validity
Ability of an indicator to measure what it is intended to measure (i.e. absence of distortions, bias or systematic errors). The most relevant biases are those related to selection of the study population and the quality of the information compiled *(23)*.

Verbal autopsy
An interview carried out with family members or caregivers of the deceased using a structured questionnaire to elicit signs and symptoms and other pertinent information that can later be used to assign a probable underlying cause of death *(150)*.

Vital event

Events that are captured through civil registration and vital statistic systems, which usually include live birth, death, fetal death, marriage, divorce, annulment of marriage, judicial separation of marriage, adoption, legitimation, recognition of parenthood, or change of name or nationality *(151)*.

Vital statistics

The systematic record of vital events such as birth, marriage, divorce, adoption, death, and cause of death to generate data and statistics *(152)*.

W

Weighting

A technique used to adjust data to reflect the known population profile. It is used to balance out any significant variance between actual and target profile. Weighted analyses of population-based surveys allow us to generalize findings to a larger or more general population. This approach aims to provide unbiased estimates of descriptive statistics or model parameters of the population of interest, which may be a general population or major population subgroup. Incorporating weights in the analyses can be crucial to achieve statistically valid, representative population-based findings in surveys and to make adjustments for sampling errors *(153)*.

Y

Years of life lost

Represents the loss of one year of life. Years of life lost are calculated from the number of deaths multiplied by a global standard life expectancy at the age at which death occurs *(154)*.

Years lived with disability

Represents the equivalent of one full year of healthy life lost due to disability or ill-health. The prevalence of years lived with disability are calculated as the prevalence of each non-fatal condition multiplied by its disability weight *(155)*.

Glossary terms by related thematic areas

Users of this glossary may wish to know which terms align with different data concepts or thematic areas. Below the glossary terms are grouped according to thematic areas. Some terms may appear in more than one thematic area.

Thematic area	Relevant terms		
Data source, collection, and collation	Data accessibility	Data type	Primary data
	Data collection level	Digital health	Response rate
	Data collection method	Geospatial data	Secondary data
	Data element	Health data	Sex
	Data information pyramid	Home-based record	Survey
	Data inputs	Measurement level	Survey data
	Data integration	Microdata	Surveillance
	Data life cycle	Population-based survey	Understandable/ simplicity
	Data source	Population census	Unit of measure
	Data standards	Preferred data sources	
Data security, protection, governance and standards	Compliance	Data steward	Mortality data
	Confidentiality	Data stewardship	National health strategic plan
	Data access control	Focal point	National statistics office
	Data anonymization	International Classification of Diseases	Notifiable conditions
	Data consent	International Health Regulations	Publicly available
	Data custodian	Master facility list	Stakeholder
	Data governance	Medical Certificate Cause of Death	Standard operating procedures
	Data privacy	Morbidity data	Sustainable Development Goals
	Data provider	Mortality coder	Transparency
	Data security		

Thematic area	Relevant terms		
Disaggregation of public health data/ indicators	Administrative area	Disaggregation	Global Health Observatory
	Age groupings recommended	Gender	Granularity
	Age specific mortality rate	Geocoding	Health inequality
	Catchment area	Geo-enablement	Health inequity
	Causes of death	Geographic accessibility	Subnational
Estimation methods	Covariates	Global health estimates	Method of aggregate estimation
	Estimation methods	Health estimates	
Health Information systems	Birth registration data	Health management information system	Registry
	Civil registration and vital statistics	Health facility census	Register
	Death registration data	Health facility survey	Registrar
	Death registration desk	Health survey	Registration form
	Disease surveillance system	Individual health record	Verbal autopsy
	Event-based surveillance	Public health surveillance system	Vital event
	Geographic information system	Raw health data	Vital statistics
	Health and demographic surveillance system	Routine data	
	Health information system	Routine health information system	

Thematic area	Relevant terms		
Indicator Terminology	Composite indicator	Metadata	Output indicator
	Core indicator	Indicator classification	Process indicator
	Domain	Indicator definition	Proxy health indicator
	Health indicator	Input indicator	Rationale
	Impact indicator	Logic model	Tracer indicator
	Indicator-based surveillance	Outcome indicator	
Methods to assess data quality	Accuracy	Data quality assurance	Punctuality
	Completeness of reporting	Internal consistency of data	Relevance
	Construct validity	Internal responsiveness	Reliability of data
	Content validity	External consistency of data	Representativeness
	Convergent validity	External responsiveness	Responsiveness
	Credibility	Feasibility	Sustainability
	Criterion validity	Gold standard	Timeliness
	Data integrity	Predictive validity	Usefulness/utility
	Data quality	Process of validation	Validity
	Data quality assessment		
Monitoring and evaluation terms	Accountability	Joint external evaluation	Results framework
	Evaluation	Monitoring	Target population
	Health impact assessment	Monitoring and evaluation Framework	
Quality of Care measurement	Care cascade	Outcome-adjusted coverage	Quality-adjusted coverage
	Effective coverage		

Thematic area	Relevant terms		
Statistical and analytical terms	Chi-square (χ^2) test	Interoperability	P value
	Confounding	Kappa statistic	Rate
	Correlation analysis	Linkage	Ratio
	Count	Mean	Record linkage
	Data	Measure	Regression
	Data harmonization	Measurement	Sample
	Data triangulation	Measurement method	Sample size
	Denominator	Median	Sampling error
	Descriptive analysis	National representative	Sensitivity analysis
	Disability adjusted life years	Negative predictive value	Statistical model
	Ecological analysis	Null hypothesis	Statistical significance
	Ecological fallacy	Numerator	Statistics
	Effect size	Odds ratio	Stratification
	Geospatial analysis	Percentage	Trend
	Healthy life expectancy at birth	Periodicity	Uncertainty measure
	Heaping of data	Population attributable fraction	Weighting
	Incidence rate	Positive Predictive Value	Years of life lost
	Imputation	Prevalence Rate	Years lived with disability
	Interaction	Processed health data	
	Intervention coverage	Proportion	

References

1. Accountability as a driver of health equity. Copenhagen: World Health Organization Regional Office for Europe; 2019 (https://iris.who.int/handle/10665/312282).
2. Data quality framework. New York: United Nations Statistics Division; 2017 (https://unstats.un.org/unsd/trade/events/2017/suzhou/presentations/Agenda%20item%2014%20-%20UNSD%20(revised).pdf, accessed 3 December 2024).
3. Second administrative level boundaries data specifications. New York: United Nations Geospatial Information Section and Statistics Division; 2021 (https://salb.un.org/sites/default/files/wysiwyg_uploads/docs_uploads/SALB_DataSpecifications.pdf, accessed 3 December 2024).
4. Diaz T, Strong KL, Cao B, Guthold R, Moran AC, Moller AB et al. A call for standardised age-disaggregated health data. Lancet. 2021;2(*7*):436–43. doi:10.1016/S2666–7568(21)00115-X.
5. Epidemiology glossary [website]. Atlanta: U.S. Centers for Disease Control and Prevention; 2024 (https://www.cdc.gov/reproductive-health/glossary/index.html, accessed 3 December 2024).
6. Selim, L. What is birth registration and why does it matter? New York: United Nations Children's Fund; 2019 (https://www.unicef.org/stories/what-birth-registration-and-why-does-it-matter, accessed 3 December 2024).
7. Tanahashi, T. Health service coverage and its evaluation. Bull World Health Organ. 1978;56(*2*):295–303 (https://iris.who.int/handle/10665/261736).
8. Marsh AD, Muzigaba M, Diaz T, Requejo J, Jackson D, Chou D et al. Effective coverage measurement in maternal, newborn, child, and adolescent health and nutrition: Progress, future prospects, and implications for quality health systems. Lancet Glob.Health. 2020;8(*5*):e730–e736. doi:10.1016/S2214–109X(20)30104–2.
9. Schuurman N, Fiedler RS, Grzybowski SCW, Grund D. Defining rational hospital catchments for non-urban areas based on travel-time. Int J Health Geogr. 2006;5(*43*). doi:10.1186/1476–072X-5–43.
10. WHO Recommendations for conducting an external inspection of a body and filling in the Medical Certificate of Cause of Death. Geneva: World Health Organization; 2022 (https://www.who.int/publications/m/item/who-recommendations-for-conducting-an-external-inspection-of-a-body-and-filling-in-the-medical-certificate-of-cause-of-death, accessed 3 December 2024).
11. Improving the quality and use of birth, death and cause-of-death information: guidance for a standards-based review of country practices. World Health Organization; 2010 (https://iris.who.int/handle/10665/44274).
12. McHugn ML. The Chi-square test of independence. Biochem Med. 2013;23: 143–49. doi:10.11613/BM.2013.018.
13. Data quality review: a toolkit for facility data quality assessment. Module 1. Framework and metrics. Geneva: World Health Organization; 2017 (https://iris.who.int/handle/10665/259224).
14. What is data compliance? New York: IBM (https://www.ibm.com/topics/data-compliance, accessed 3 December 2024).
15. The handbook on management and organization of national statistical systems. New York: United Nations Department of Economic and Social Affairs. 2022; (https://unstats.un.org/capacity-development/handbook/html/topic.htm#t=Handbook%2FCover%2FCover.htm, accessed 3 December 2024).
16. Adler S. What is the confidentiality definition in healthcare? [website] HIPAA Journal; 2024 (https://www.hipaajournal.com/confidentiality-definition-in-healthcare/, accessed 3 December 2024).

17. Porta M. A dictionary of epidemiology (6 ed.). New York: Oxford University Press; 2014 (https://www.oxfordreference.com/display/10.1093/acref/9780199976720.001.0001/acref-9780199976720-e-374?rskey=GUqlHJ&result=430, accessed 3 December 2024).
18. Strauss ME, Smith GT. Construct validity: advances in theory and methodology. Annu Rev Clin Psychol. 2009;5:1–25. doi:10.1146/annurev.clinpsy.032408.153639.
19. de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in medicine: A practical guide (Practical Guides to Biostatistics and Epidemiology). Cambridge: Cambridge University Press; 2011.
20. Krabbe PFM. The measurement of health and health status. Academic Press; 2017 (https://www.sciencedirect.com/science/article/abs/pii/B9780128015049000064, accessed 3 December 2024).
21. 2018 Global reference list of 100 core health indicators (plus health-related SDGs). Geneva: World Health Organization; 2018 (https://iris.who.int/handle/10665/259951).
22. Babbie E. The practice of social research. Boston: Cengage Learning; 2021.
23. Health indicators. Conceptual and operational considerations. Washington DC: Pan American Health Organization; 2018 (https://iris.paho.org/handle/10665.2/49056).
24. Stevens GA, Alkema L, Black RE, Boerma JT, Collins GS, Ezzati M. Guidelines for accurate and transparent health estimates reporting: the GATHER statement. Lancet. 2016;388(*10062*):E19–E23. doi:10.1016/S0140–6736(16)30388–9.
25. Richiardi L. A dictionary of epidemiology (6 ed.). New York: Oxford University Press; 2014.
26. MacFeely S. In search of the data revolution: Has the official statistics paradigm shifted?. Stat J IAOS. 2020;36(*4*):1075–94. doi:10.3233/SJI-200662.
27. Kickbusch I, Piselli D, Agrawal A, Balicer R, Banner O, Adelhardt M. The Lancet and Financial Times Commission on governing health futures 2030: growing up in a digital world. Lancet. 2021; 398:1727–1776. doi:10.1016/S0140–6736(21)01824–9.
28. Data strategy of the Secretary-General for Action by Everyone, Everywhere with Insight, Impact and Integrity 2020–22. New York: United Nations; 2020 (https://www.un.org/en/content/datastrategy/images/pdf/UN_SG_Data-Strategy.pdf, accessed 3 December 2024).
29. Glossary terms [website]. Herziliyah Pituach: Gigaspaces (https://www.gigaspaces.com/data-terms, accessed 3 December 2024).
30. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). Brussels: European Union; 2016 (https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679, accessed 3 December 2024).
31. World Health Organization (WHO) Personal Data Protection Policy. Geneva: World Health Organization; 2024 (https://cdn.who.int/media/docs/default-source/world-health-data-platform/who_personal_data_protection_policy.pdf?sfvrsn=f7089a28_3&download=true, accessed 3 December 2024).
32. Hjollund NHI, Valderas JM, Kyte D, Calvert MJ. Health Data Processes: A framework for analyzing and discussing efficient use and reuse of health data with a focus on patient-reported outcome measures. J Med Internet Res. 2019;21(*5*):e12412. doi:10.2196/12412.
33. Metadata format [website]. Open SDG. (https://open-sdg.readthedocs.io/en/latest/metadata-format/, accessed 3 December 2024).
34. Rights, Roles and Responsibilities. Data Custodians [website]. Canberra: Australian Government (https://toolkit.data.gov.au/data-integration/roles-and-responsibilities/data-custodians.html, accessed 3 December 2024).
35. NIST Privacy framework: A tool for improving privacy through enterprise risk management. Gaithersburg: National Institute of Standards and Technology U.S. Department of Commerce; 2020 (https://www.nist.gov/system/files/documents/2020/01/16/NIST%20Privacy%20Framework_V1.0.pdf, accessed 3 December 2024).

36. Rosenbaum S. Data governance and stewardship: Designing data stewardship entities and advancing data access. Health Serv Res. 2010;45(5 Pt 2):1442–1455. doi:10.1111/j.1475–6773.2010.01140.x.

37. Data harmonization [website]. Ann Arbor: Data sharing for demographic research (https://www.icpsr.umich.edu/web/pages/DSDR/harmonization.html#:~:text=Data%20harmonization%20refers%20to%20all,of%20data%20from%20different%20studies, accessed 3 December 2024).

38. Using information for action: A manual for health workers at facility level. Pretoria: The EQUITY Project University of Western Cape (https://pdf.usaid.gov/pdf_docs/Pnacx654.pdf, accessed 8 February 2024).

39. Data integration in healthcare: Guide and best practices [website]. Atlanta: KMS Healthcare; 2024 (https://kms-healthcare.com/blog/data-integration-in-healthcare/, accessed 3 December 2024).

40. Leventer-Roberts M, Balicer R. Data integration in health care. In: Amelung V, Stein V, Suter E, Goodwin N, Nolte E, Balicer R, editors. Handbook Integrated Care. Berlin: Springer; 2021.

41. Integrity [website]. Global Health eLearning Center (https://www.globalhealthlearning.org/taxonomy/term/992, accessed 3 December 2024).

42. Guidance document on good in vitro method practices (GIVIMP). Paris: Organisation for Economic Co-operation and Development Publishing; 2018 (https://www.oecd-ilibrary.org/docserver/9789264304796-en.pdf, accessed 3 December 2024).

43. Data for Better Lives. World Development Report 2021. Washington DC: World Bank; 2021 (https://www.worldbank.org/en/publication/wdr2021, accessed 3 December 2024).

44. Data management competency framework. Manila: World Health Organization Regional Office for the Western Pacific; 2023 (https://iris.who.int/handle/10665/367502).

45. Data privacy at CERN. Geneva: CERN (https://privacy.web.cern.ch/terminology-page, accessed 3 December 2024).

46. Ensuring the integrity, accessibility, and stewardship of research data in the digital age. Washington DC: The National Academies Press; 2009 (https://www.ncbi.nlm.nih.gov/books/NBK215264/, accessed 3 December 2024).

47. Data standards. ID4D Practitioner's Guide. Washington DC: World Bank (https://id4d.worldbank.org/guide/data-standards, accessed 3 December 2024).

48. Aspden P, Corrigan JM, Wolcott J, Erickson SM. Patient safety: Achieving a new standard for care. Washington DC: National Academies Press; 2004 (https://pubmed.ncbi.nlm.nih.gov/25009854/, accessed 3 December 2024).

49. Data quality assurance. Module 1. Framework and metrics. Geneva: World Health Organization; 2022 (https://iris.who.int/handle/10665/366086, accessed 3 December 2024).

50. Framework and standards for country health information systems. Geneva: World Health Organization; 2012 (https://iris.who.int/handle/10665/43872).

51. Toolkit for analysis and use of routine health facility data. Integrated health services analysis: national level. Geneva: World Health Organization; 2023 (https://iris.who.int/handle/10665/369381).

52. What is data stewardship? Its importance, benefits, programs and more [website]. Simplilearn (https://www.simplilearn.com/what-is-data-stewardship-article, accessed 3 December 2024).

53. Defining data stewardship: A collaborative global discourse. New York: United Nations Department of Economic and Social Affairs; 2024 (https://unstats.un.org/UNSDWebsite/statcom/session_55/side-events/UNSC55-side-event-Defining-Data-Stewardship-A-Collaborative-Global-Discourse.pdf, accessed 3 December 2024).

54. An introduction to triangulation. Geneva: UNAIDS; 2013 (https://www.unaids.org/sites/default/files/sub_landing/files/10_4-Intro-to-triangulation-MEF.pdf, accessed 3 December 2024).

55. Rutherford GW, McFarland W, Spindler H, White K, Patel SV, Aberle-Grasse J et al. Public health triangulation: approach and application to synthesizing data to understand national and local HIV epidemics. BMC Public Health. 2010;10:447. doi:10.1186/1471–2458–10–447.

56. Guetterman TC. Basics of statistics for primary care research. Fam Med Community Health. 2019;7(*2*):e000067. doi:10.1136/fmch-2018–000067.

57. Lesko CR, Fox MP, Edwards JK. A framework for descriptive epidemiology. Am J Epidemiol. 2022;191(*12*):2063–70. doi:10.1093/aje/kwac115.

58. Digital implementation investment guide (DIIG): Integrating digital interventions into health programmes. Geneva: World Health Organization; 2020 (https://iris.who.int/handle/10665/334306).

59. Disability-adjusted life years (DALYs). The Global Health Observatory [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/158#:~:text=Definition%3A-,One%20DALY%20represents%20the%20loss%20of%20the%20equivalent%20of%20one,health%20condition%20in%20a%20population, accessed 3 December 2024).

60. Communicable disease surveillance systems [website]. Cairo: World Health Organization Regional Office for the Eastern Mediterranean (https://www.emro.who.int/health-topics/health-information-systems/communicable-disease-surveillance-systems.html, 3 December 2024).

61. Ng M, Fullman N, Dieleman JL, Flaxman AD, Murray CJ, Lim SS. Effective coverage: a metric for monitoring Universal Health Coverage. PLoS Med. 2014;11(*9*):e1001730. doi:10.1371/journal.pmed.1001730.

62. Karim A, de Savigny D. Effective coverage in health systems: Evolution of a concept. Diseases. 2023 Feb 22;11(*1*):35. doi:10.3390/diseases11010035.

63. Effect size. ScienceDirect [website]. Amsterdam: Elsevier; 2024 (https://www.sciencedirect.com/topics/nursing-and-health-professions/effect-size#:~:text=Effect%20size%20refers%20to%20the,larger%20sample%20size%20for%20detection, accessed 3 December 2024).

64. A guide to establishing event-based surveillance. Manila: World Health Organization Regional Office for the Western Pacific; 2008 (https://iris.who.int/handle/10665/207737).

65. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness a critical review and recommendations. J Clin Epidemiol. 2000; 53(*5*):459–68. doi:10.1016/S0895–4356(99)00206–1.

66. Communicable disease surveillance and response systems: Guide to monitoring and evaluating. World Health Organization; 2006 (https://iris.who.int/handle/10665/69331).

67. Country portal [website]. Geneva: World Health Organization; 2024 (https://data.who.int/about/data/whdh/country-portal, accessed 3 December 2024).

68. Gender and health [website]. Geneva: World Health Organization (https://www.who.int/health-topics/gender, accessed 3 December 2024).

69. Introduction to finding places on a map [website]. Aylesbury: Esri (https://pro.arcgis.com/en/pro-app/latest/help/data/geocoding/introduction-to-finding-places-on-a-map.htm, accessed 3 December 2024).

70. Geo-enabled microplanning handbook. WHO-UNICEF COVAX GIS Working Group (https://knowledge.unicef.org/resource/geo-enabled-microplanning-handbook, accessed 9 December 2024).

71. AccessMod 5. Supporting Universal Health Coverage by modelling physical accessibility to health care [website]. Geneva: AccessMod (https://www.accessmod.org/)

72. What is GIS? Esri (https://www.esri.com/en-us/what-is-gis/overview#:~:text=Geographic%20information%20system(s)%2C,what%20things%20are%20like%20there, accessed 3 December 2024).

73. WHO GIS Centre for Health [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/GIS, accessed 3 December 2024).

74. Boots B. Spatial Pattern, Analysis of. In: Smelser NJ, Baltes PB, editors. International Encyclopedia of the Social & Behavioral Sciences. Pergamon; 2001: 14818–14822.

75. About the Observatory. The Global Health Observatory [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/gho/info/about-the-observatory, accessed 3 December 2024).

76. Global health estimates. Geneva: World Health Organization; 2024 (https://www.who.int/data/global-health-estimates, accessed 3 December 2024).

77. Gold standard. Glossary [website]. Manchester: National Institute for Health and Care Excellence (https://www.nice.org.uk/glossary?letter=g#:~:text=Gold%20standard,for%20or%20treat%20a%20disease, 3 December 2024).

78. Data Granularity. Glossary [website]. Redwood City: C3.AI (https://c3.ai/glossary/features/data-granularity/, accessed 3 December 2024).

79. Cunningham SA, Shaikh NI, Nhacolo A, Raghunathan PL, Kotloff K, Naser AM et al. Health and Demographic Surveillance Systems Within the Child Health and Mortality Prevention Surveillance Network. Clin Infect Dis. 2019; 69(Suppl 4):S274–79. doi:10.1093/cid/ciz1162.

80. Introduction to basic data analysis and interpretation for health programs. A training tool kit module 1: Data analysis key concepts [website]. Chapel Hill: MEASURE Evaluation. United States Agency for International Development (https://www.measureevaluation.org/resources/training/capacity-building-resources/basic-data-analysis-for-health-programs.html, accessed 3 December 2024).

81. Greenwell F, Salentine S. Health information system strengthening: standards and best practices for data sources. Chapel Hill: MEASURE Evaluation. United States Agency for International Development; 2018 (https://www.measureevaluation.org/resources/publications/tr-17–225.html, 3 December 2024).

82. Vohra S, Cave B, Viliani F, Harris-Roxas BF, Bhatia R. New international consensus on health impact assessment. The Lancet. 2010; 376(*9751*):1464–1465. doi:10.1016/S0140–6736(10)61991–5.

83. National health inequality monitoring: a step-by-step manual. Geneva: World Health Organization; 2017 (https://iris.who.int/handle/10665/255652).

84. Guidelines for the development of health management information systems. Manila: World Health Organization Regional Office for the Western Pacific; 1993 (https://iris.who.int/handle/10665/207002).

85. Developing health management information systems: a practical guide for developing countries. Manila: World Health Organization Regional Office for the Western Pacific; 2004 (https://iris.who.int/handle/10665/207050).

86. Terms related to HIS. HIS Strengthening Resource Center [website]. Chapel Hill: MEASURE Evaluation. United States Agency for International Development (https://www.measureevaluation.org/his-strengthening-resource-center/his-definitions/terms-related-to-his.html, accessed 3 December 2024).

87. Tulchinsky TH, Varavikova EA. Measuring, monitoring, and evaluating the health of a population. The New Public Health (Third Edition). Amsterdam: Elsevier; 2014 (https://pmc.ncbi.nlm.nih.gov/articles/PMC7171905/, accessed 3 December 2024).

88. Healthy life expectancy (HALE) at birth. The Global Health Observatory [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/66#:~:text=Definition%3A,to%20disease%20and%2For%20injury, accessed 3 December 2024).

89. Jaeschke L, Becher M, Velásquez IM, Ahrens W, Bächle C, Baurecht H et al. The bias from heaping on risk estimation: effect of age at diagnosis of hypertension on risk of subsequent cardiovascular comorbidities. Ann Epidemiol. 2022; 74:84–96. doi:10.1016/j.annepidem.2022.07.012.

90. WHO recommendations on home-based records for maternal, newborn and child health. Geneva: World Health Organization; 2018 (https://iris.who.int/handle/10665/274277).

91. Introduction to data imputation [website]. Simplilearn; 2023 (https://www.simplilearn.com/data-imputation-article#:~:text=Data%20imputation%20is%20a%20method,data%20with%20a%20different%20value, accessed 3 December 2024).

92. Scientific and technical review of the traditional knowledge indicators and their suggested links with the headline, component and complementary indicators of the monitoring framework for the Kunming-Montreal Global Biodiversity Framework. Convention on Biological Diversity. Nairboi: UN Environment Programme; 2024 (https://www.cbd.int/doc/c/283e/eb7c/6a953a1e098b6c46e0f3be8d/sbstta-26-inf-11-en.pdf, accessed 3 December 2024).

93. Country has national electronic health record (EHR). Global Health Observatory [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/4791, accessed 3 December 2024).

94. Internal consistency reliability: definition, examples. Statistics How To (https://www.statisticshowto.com/internal-consistency/, accessed 3 December 2024).

95. International Health Regulations (*2005*), 3rd ed. Geneva: World Health Organization; 2005 (https://iris.who.int/handle/10665/246107).

96. Monitoring the building blocks of health systems: A handbook of indicators and their measurement strategies. Geneva: World Health Organization; 2010 (https://iris.who.int/handle/10665/258734).

97. Joint External Evaluation (JEE) [website]. Geneva: World Health Organization (https://www.who.int/emergencies/operations/international-health-regulations-monitoring-evaluation-framework/joint-external-evaluations, accessed 3 December 2024).

98. Fleiss JL, Cohen J, Everitt BS. Large sample standard errors of kappa and weighted kappa. Psychological Bulletin. 1969;72(*5*):323–327. doi:10.1037/h0028106.

99. Harron K. Data linkage in medical research. BMJ Med. 2022;1:e000087. doi:10.1136/bmjmed-2021–000087.

100. Surveillance, monitoring, and evaluation of malaria programs [online course]. Chapel Hill: MEASURE Evaluation. United States Agency for International Development; 2020 (https://www.measureevaluation.org/resources/publications/ms-20–184/at_download/document, accessed 3 December 2024).

101. Demystifying the M&E framework: A guide for effective evaluation [website]. Evalcommunity (https://www.evalcommunity.com/career-center/me-framework/, accessed 3 December 2024).

102. Geolocated health facilities data initiative [website]. Geneva: World Health Organization; 2022 (https://www.who.int/news-room/questions-and-answers/item/global-health-facilities-database, 3 December 2024).

103. Mean, median and mode [website]. Newcastle: Newcastle University (https://www.ncl.ac.uk/webtemplate/ask-assets/external/maths-resources/statistics/descriptive-statistics/mean-median-and-mode.html, accessed 3 December 2024).

104. Sustainable Development Goals [website]. Geneva: World Health Organization (https://www.who.int/europe/about-us/our-work/sustainable-development-goals, accessed 3 December 2024).

105. SDMX content-oriented guidelines: Metadata common vocabulary. Statistical Data and Metadata eXchange; 2006 (https://sdmx.org/wp-content/uploads/Content_04_Draft_Guidelines_Metadata_Common_Vocabulary-MARCH-2006–1.pdf, accessed 3 December 2024).

106. Guidelines for the global metadata structure definition for Sustainable Development Goals indicators. New York: United Nations Statistical Division; 2023 (https://unstats.un.org/sdgs/files/SDG_MSD_Guidelines.pdf, accessed 3 April 2024).

107. What do we mean by microdata? Washington DC: World Bank (https://datahelpdesk.worldbank.org/knowledgebase/articles/228873-what-do-we-mean-by-microdata, accessed 3 December 2024).

108. Noncommunicable disease surveillance, monitoring and reporting. World Health Organization (https://www.who.int/teams/noncommunicable-diseases/surveillance/systems-tools/patient-programme-monitoring, accessed 3 December 2024).

109. Hernandez JBR, Kim PY. Epidemiology morbidity and mortality. Florida: StatPearls Publishing; 2022 (https://pubmed.ncbi.nlm.nih.gov/31613448/, accessed 3 December 2024).

110. Strategizing national health in the 21st century: a handbook. Geneva: World Health Organization; 2016 (https://iris.who.int/handle/10665/250221).

111. Recommendation of the Council on Good Statistical Practice [website]. Paris: Organisation for Economic Co-operation and Development Publishing (https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0417, accessed 3 December 2024).

112. Methodological principles of nationally representative surveys as a platform for global surveillance of antimicrobial resistance in human bloodstream infections. Geneva: World Health Organization; 2023 (https://iris.who.int/handle/10665/366150).

113. Altman DG, Bland JM. Statistics Notes: Diagnostic tests 2: predictive values. BMJ. 1994;209:102. doi:10.1136/bmj.309.6947.102.

114. Notifiable disease. Atlanta: National Center for Health Statistics; 2024 (https://www.cdc.gov/nchs/hus/sources-definitions/notifiable-disease.htm, accessed 3 December 2024).

115. Moore DS, McCabe GP. Introduction to the Practice of Statistics (4 ed.). New York: W.H. Freeman and Co; 2003.

116. CDC approach to programme evaluation. Atlanta: U.S. Centers for Disease Control and Prevention; 2021 (https://www.cdc.gov/evaluation/php/about/index.html, accessed 7 May 2024).

117. Cambridge Dictionary. Cambridge: Cambridge University Press (https://dictionary.cambridge.org/dictionary/, accessed 3 December 2024).

118. Population attributable fraction. The Global Health Observatory [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/1287, accessed 3 December 2024).

119. Principles and recommendations for population and housing censuses. New York: United Nations Department of Economic and Social Affairs; 2017 (https://unstats.un.org/unsd/publication/seriesM/Series_M67rev3en.pdf, accessed 3 December 2024).

120. Lesson 3: Measures of risk. Atlanta: US Centers for Disease Control and Prevention. 2012; (https://archive.cdc.gov/www_cdc_gov/csels/dsepd/ss1978/lesson3/section2.html, 3 December 2024).

121. Data availability. A visual summary [website]. Geneva: World Health Organization (https://www.who.int/data/gho/data-availability-a-visual-summary, accessed 3 December 2024).

122. An introduction to indicators. Geneva: UNAIDS; 2010 (https://www.unaids.org/sites/default/files/sub_landing/files/8_2-Intro-to-IndicatorsFMEF.pdf, accessed 3 December 2024).

123. Introduction to public health surveillance. Atlanta: U.S. Centers for Disease Control; 2024 (https://www.cdc.gov/training-publichealth101/php/training/introduction-to-public-health-surveillance.html, accessed 3 December 2024).

124. Regulatory definitions of publicly available, public domain, and published overview. Tony Brook: Stony Brook University (https://www.stonybrook.edu/commcms/export-controls/Export-Control-Regulations/Publicly-Available.php, accessed 3 December 2024).

125. Ensor T, Cooper S. Overcoming barriers to health service access: influencing the demand side. Health Policy Plan. 2004;19(*2*):69–79. doi:10.1093/heapol/czh009.

126. Amouzou A, Leslie HH, Ram M, Fox M, Jiwani SS, Requejo J et al. Advances in the measurement of coverage for RMNCH and nutrition: from contact to effective coverage. BMJ Glob Health. 2019; 4:i114–i124. doi:10.1136/bmjgh-2018–001297.

127. Schirnding YV. Health in sustainable development planning: The role of indicators. Geneva: World Health Organization; 2002 (https://iris.who.int/handle/10665/67391).

128. Winkler WE. The state of record linkage and current research problems. Washington, DC: U. S. Bureau of the Census (https://www.census.gov/content/dam/Census/library/working-papers/1999/adrm/rr99–04.pdf, accessed 3 December 2024).

129. Principles and recommendations for a vital statistics system (revision 3). New York: United Nations Department of Economic and Social Affairs; 2014 (https://unstats.un.org/unsd/demographic/standmeth/principles/m19rev3en.pdf, accessed 3 December 2024).

130. Brooke EM. The current and future use of registers in health information systems. Geneva: World Health Organization; 1974 (https://iris.who.int/handle/10665/36936).

131. Handbook on civil registration and vital statistics systems management, operation and maintenance. New York: United Nations Department of Economic and Social Affairs; 2021 (https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/crvs/crvs-mgt-E.pdf, accessed 3 December 2024).
132. What is a response rate? Hoboken: NICE (https://www.nice.com/glossary/what-is-a-response-rate, accessed 3 December 2024).
133. Estrada E, Ferrer E, Pardo A. Statistics for evaluating pre-post change: Relation between change in the distribution center and change in the individual scores. Front Psychol. 2019;9:2696. doi:10.3389/fpsyg.2018.02696.
134. WHO Toolkit for routine health information systems data. Geneva: World Health Organization https://www.who.int/data/data-collection-tools/health-service-data/toolkit-for-routine-health-information-system-data/modules, accessed 3 December 2024).
135. RHIS Curriculum. Chapel Hill: MEASURE Evaluation. United States Agency for International Development (https://www.measureevaluation.org/our-work/routine-health-information-systems/rhis-curriculum, accessed 3 December 2024).
136. Banerjee A, Chaudhury S. Statistics without tears: Populations and samples. Ind Psychiatry J. 2010;19(*10*):60–65. doi:10.4103/0972–6748.77642.
137. Glossary. Expert Group on National Quality Assurance Frameworks [website]. United Nations Statistics Division; 2012 (https://unstats.un.org/unsd/dnss/docs-nqaf/NQAF%20GLOSSARY.pdf, accessed 3 December 2024).
138. Saltelli A, Tarantola S, Campolongo F, Ratto M. Sensitivity Analysis in Practice: A guide to assessing scientific models. Oxford: John Wiley & Sons Ltd; 2004 (https://www.andreasaltelli.eu/file/repository/SALTELLI_2004_Sensitivity_Analysis_in_Practice.pdf, accessed 3 December 2024).
139. Miller D, Oliver M. Engaging stakeholders for project success [website]. Project Management Institute; 2015 (https://www.pmi.org/learning/library/engaging-stakeholders-project-success-11199, 3 December 2024).
140. Hashemi-Pour C. Standard operating procedure (SOP) [website]. London: Techtarget (https://www.techtarget.com/searchbusinessanalytics/definition/standard-operating-procedure-SOP, accessed 3 December 2024).
141. Marimuthi R, Shivappriya SN, Saroja MN. A study of machine learning algorithms used for detecting cognitive disorders associated with dyslexia. In: Jude HD, editor. Handbook of decision support systems for neurological disorders. Cambridge: Academic Press. 2021:245–62. doi:10.1016/B978–0–12–822271–3.00008–6.
142. Tenny S, Abdelgawad I. Statistical significance. Florida: StatPearls Publishing; 2023 (https://www.ncbi.nlm.nih.gov/books/NBK459346/, accessed 19 July 2024).
143. MacFeely S. You say you want a [data] revolution: A proposal to use unofficial statistics for the SDG Global Indicator Framework. Stat J IAOS. 2019;35(*3*):309–27. doi:10.3233/SJI-180486.
144. Updated guidelines for evaluating public health surveillance systems: Recommendations from the Guidelines Working Group. MMWR. 2001;50(RR13):1–35. Atlanta: US Centers for Disease Control and Prevention; 2001 (https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5013a1.htm, accessed 3 December 2024).
145. Terminology on statistics metadata. Geneva: United Nationals Statistical Commission and Economic Commission for Europe; 2000 (https://unece.org/DAM/stats/publications/53metadaterminology.pdf, accessed 3 December 2024).
146. Statistics Canada Quality Guidelines. Ottawa: Statistics Canada; 2009 (https://www150.statcan.gc.ca/n1/pub/12–539-x/12–539-x2009001-eng.pdf, accessed 3 December 2024).
147. The right to be informed (transparency) (Article 13 & 14 GDPR). Dublin: Data Protection Commission (https://www.dataprotection.ie/en/individuals/know-your-rights/right-be-informed-transparency-article-13–14-gdpr, accessed 3 December 2024).
148. Trend [definition]. New York: Statista (https://www.statista.com/statistics-glossary/definition/425/trend/, accessed 3 December 2024).

149. Uncertainty and how we measure it for our surveys. London: Office for National Statistics; 2018 (https://www.ons.gov.uk/methodology/methodologytopicsandstatisticalconcepts/uncertaintyandhowwemeasureit, 3 December 2024).

150. Cause of Death. International Classification of Diseases (ICD-11). Geneva: World Health Organization (https://www.who.int/standards/classifications/classification-of-diseases/cause-of-death, accessed 3 December 2024).

151. Guidelines on the legislative framework for civil registration, vital statistics and identity management systems. New York: United Nations Department of Economic and Social Affairs; 2023 (https://desapublications.un.org/publications/guidelines-legislative-framework-civil-registration-vital-statistics-and-identity, accessed 3 December 2024).

152. Civil registration and vital statistics [website]. Geneva: World Health Organization; 2024 (https://www.who.int/news-room/fact-sheets/detail/civil-registration-and-vital-statistics, accessed 3 December 2024).

153. Si Yajuan, Lee S, Heeringa SG. Population Weighting in Statistical Analysis. Jama Intern.Med. 2024;184(*1*):98–99. doi:10.1001/jamainternmed.2023.6300.

154. Years of life lost from mortality (YLL). Global Health Observatory [website]. Geneva: World Health Organization; 2024 (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/159, accessed 3 December 2024).

155. Years of healthy life lost due to disability (YLD). Global Health Observatory [website] Geneva: World Health Organization (https://www.who.int/data/gho/indicator-metadata-registry/imr-details/160, accessed 3 December 2024).

For further information, please contact:
Data Hub and Spoke Collaborative Secretariat
World Health Organization
Email: datagovernance@who.int