Since 2016 the World Health Statistics series has served as WHO’s annual report on the health-related Sustainable Development Goals (SDGs). The effective monitoring of SDG indicators requires comprehensive national health information strategies based on the use of data from sources such as civil registration and vital statistics systems, household and other population-based surveys, routine health-facility reporting systems and health-facility surveys, administrative data systems and surveillance systems. Some indicators also rely on non-health-sector data sources.

Making sense of the often complex available data on health indicators can be highly challenging. Health data derived from health information systems, including health-facility records, surveys or vital statistics, may not be representative of the entire population of a country and in some cases may not even be accurate. Comparisons between populations or over time can also be complicated by differences in data definitions and/or measurement methods. Although some countries may have multiple sources of data for the same year, it is more usual for data not to be available for every population or year. For example, measurement frequency for data collected through household surveys is typically every 3–5 years. This means that the years for which data are available differ by country. To overcome these and other issues and allow for comparisons to be made across countries and over time, analysts develop mathematical and statistical models with the aim of producing unbiased estimates that are representative and comparable.

In World health statistics 2017, more than 50 health-related SDG indicators were identified. Currently, sufficient monitoring data are available for 36 indicators and these data are presented in Annexes A and B of the current report, as well as online in the WHO Global Health Observatory (www.who.int/gho/en). For most indicators, comparable estimates are reported if they are available. Such data have been generated using a database of primary data and a mathematical or statistical model, followed by consultation with the relevant WHO Member State. In these cases, the database of primary data used to derive the estimates is available online, together with other documentation required by the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER).

For other indicators, the most recent observation from a database of primary data is reported. Primary data is

---

Fig. 1.1
Schematic overview of WHO data compilation, processing, analysis, consultation and reporting

1 DATA COLLECTION
WHO collects data from a variety of sources through official requests or accessing publicly available data.

2 DATA COMPILATION
While some primary data need to be compiled, processed and validated, other primary data are forwarded for publication.

3 COMPARABLE ESTIMATES
A statistical or mathematical model is used to calculate comparable estimates.

4 CONSULTATION
After initial statistics are obtained, Member States are given an opportunity to comment on methods or provide new primary data.

5 REPORTING
Data and methods are reported in line with GATHER.

6 PUBLICATION AND DISSEMINATION
WHO publishes its official statistics through its flagship products such as the World Health Statistics series.

Key informant data may be validated before official publication.

Note: Data on emergencies does not pass through the country consultation and reporting process in most cases.
An umbrella term that includes both raw data (measures derived from primary data collection with no adjustments or corrections) and processed data (calculated from raw data).\(^1\) Processing raw health data can include cleaning data by removing implausible values, calculating an indicator with an algorithm or adjusting a statistic for bias. In some, but not all, cases these data have been consulted upon with each respective Member State.

Although most data series reported in World Health Statistics are either compilations of primary data or comparable estimates, there are some data series which do not clearly fit into either of these categories. Typically these are data series compiled using the results of surveys of key informants, such as government officials, in countries. Such data series may reflect primary data known to the informant, estimates known to the informant, or the opinion of the informant regarding the local situation. In order to label such data in the current report, a third data category – other data – is used.

A schematic overview of the compilation and processing of primary data, calculation of comparable estimates, consultation with Member States and publication in the World Health Statistics and other World Health Organization data products is provided in Fig. 1.1.

In World health statistics 2018, each data series has for the first time been labelled as “comparable estimates”, most recent “primary data” or “other data” to clearly indicate the category to which it belongs. The features of each of these three types of data series are outlined in Table 1.1. These data labels can be used by readers of this report to guide interpretation of the data presented and to inform further investigation on data sources by topic. Users of comparable estimates should interrogate the availability and quality of the underlying data used to generate the estimates, and should take into account uncertainty intervals (available online at the WHO Global Health Observatory). Users of primary data should assess whether the data are comparable, taking into account the inclusion/exclusion criteria for the database, whether adjustments were made to improve comparability and the year of data collection.

In this regard, attention should be given to the footnotes on country statistics provided in Annex B. Finally, users of statistics which are labelled as other data should be aware that primary data may not be available, and that data are often not comparable across countries.

In addition to the importance of understanding these different types of information at the global level to inform interpretation and policy dialogue, the reviewing of data sources and data availability at country level can also help to define the scope of ongoing and future health information strategies. In particular, any gaps in data collection can be identified and solutions prioritized to support the development of informed national health strategic plans.