

THELANCET-D-19-04428R1

S0140-6736(19)32523-1

Doctopic: Review and Opinion



Monitoring the health-related Sustainable Development Goals: lessons learned and recommendations for improved measurement

Samira Asma, Rafael Lozano, Somnath Chatterji, Soumya Swaminathan, Maria de Fátima Marinho, Naoko Yamamoto, Elena Varavikova, Awoke Misganaw, Michael Ryan, Lalit Dandona, Christopher J L Murray

Background

The UN General Assembly launched the Sustainable Development Goals (SDGs) in September, 2015.¹ The original global SDG framework included 17 goals, 169 targets, and 232 unique indicators.² Of these, 12 goals, 33 targets, and 57 indicators have been identified as health-related SDGs (HRSDGs),^{3–6} that is, pertaining to health outcomes, health services, and well-established environmental, occupational, behavioural, and metabolic risks. The scope of health in the SDGs is much broader than in the Millennium Development Goals, spanning from maternal and child health and infectious diseases to non-communicable diseases, injuries, risk factors, and health-system functions. Regular monitoring of the HRSDGs is important for fostering a shared notion of accountability for results, identifying important gaps in resources and rates of progress, and taking into account emerging challenges that can influence the trajectory of progress. Regular monitoring and accountability will be essential to sustain policy focus and funding for the broad and complex HRSDG agenda.

100 countries have published SDG monitoring reports since 2015 and, in 2019, 38 more countries indicated intentions to report.^{7,8} In addition to these government-led efforts, several international groups report on the HRSDGs: WHO in 2016, 2017, 2018, and 2019;^{5,9–11} the World Bank in 2017 and 2018;^{12,13} the Global Burden of Disease (GBD) collaboration in 2016, 2017, and 2018;^{6,14,15} the Sustainable Development Solutions Network in 2016, 2017, and 2018;^{16–18} and Our World in Data's SDG dashboard starting in 2017–18.¹⁹ These reporting efforts on the HRSDGs differ in the number of indicators, countries, and years covered. Where reports overlap for the same indicator, country, and year, correlation coefficients of the estimates vary widely. For example, WHO and GBD's most recent reports have correlation coefficients varying from 0·94 for under-five mortality (indicator 3.2.1) to 0·43 for road traffic mortality (indicator 3.6.1).^{10,15} Poor correlations for some indicators across reporting efforts highlight inconsistencies that emerge from using different HRSDG definitions, data sources, data processing, and data synthesis approaches.

In this Viewpoint, we examine why HRSDG results can differ so much across these empirical monitoring efforts and make recommendations on moving towards more standardised, universal assessments.

Defining HRSDG indicators

Beyond SDG 3—the SDG dedicated to ensuring healthy lives and promoting wellbeing for all—the global indicator

framework includes many indicators that directly measure health outcomes (eg, indicator 3.1.1; maternal mortality ratio) or determinants of health (eg, indicator 6.1.1; proportion of population using safely managed drinking water services). Yet, defining what constitutes a HRSDG inevitably involves a judgment about the relative proximal or distal relationship of an indicator to health, as well as its scope within the direct or stewardship functions of the health system in delivering health interventions or reducing risks to health.

Drawing from previous monitoring efforts by WHO and the GBD collaboration,^{11,15,20,21} we identified 57 HRSDG indicators that (1) either are or directly relate to health outcomes and risks, health services and interventions, or health system needs and capacities (eg, indicator 3.d.1; international health regulations capacity and health emergency preparedness, indicator 17.19.2c; well certified deaths); and (2) have been established as health priorities by countries via international agendas or agencies (appendix pp 1–4). These criteria lead to a larger set of HRSDGs than had been previously reported by UN agencies and the GBD collaboration, with the latter identifying 52 HRSDGs in its 2018 analysis.¹⁵ Continuously updating the number of indicators included within the HRSDGs—and thus their scope—is far from ideal, but we view this revisiting process as crucial for better aligning HRSDG monitoring with recent global initiatives (eg, WHO's Global Action Plan on the Sustainable Development Goals, which was initially launched in October, 2018).²⁰ Further, with its annual indicator refinements and updates via the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs) and UN Statistical Commission, as well as its upcoming comprehensive reviews in 2020 and 2025, the global SDG indicator framework remains a dynamic development agenda.

How each HRSDG is operationalised—that is, how indicators from the global framework are translated to meta-data and then to quantifiable measures at the country level—can contribute to differences in monitoring efforts. For some HRSDGs, particularly those for which data are incomplete or sparse in many countries, the proxy measures selected or modifications implemented could underlie differences in HRSDG reporting, in addition to data input and analyses. Continued collaboration to bring proxy measures closer to indicator definitions and their meta-data is important. Doing so also hinges upon heightened investments in data collection systems as well as testing which proxy

WHO, Geneva, Switzerland (S Asma DDS, S Chatterji MD, S Swaminathan MD, N Yamamoto MD, M Ryan MD); Department of Health Metrics Sciences, Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, USA (Prof R Lozano MD, A Misganaw PhD, Prof L Dandona MD, Prof C J L Murray DPhil); Institute of Advanced Studies, University of São Paulo, São Paulo, Brazil (M de Fátima Marinho PhD); Federal Research Public Health Organization and Information Institute, Ministry of Health, Moscow, Russia (E Varavikova PhD); Ethiopian Public Health Institute, Addis Ababa, Ethiopia (A Misganaw); and Public Health Foundation of India, Gurugram, National Capital Region, India (Prof L Dandona)

Correspondence to:

Dr Samira Asma, World Health Organization, 1211, Switzerland
asmas@who.int

or

Prof Rafael Lozano, University of Washington, 98121, USA
rlozano@uw.edu

See Online for appendix

indicators are most closely predictive of recent HRSDG patterns and trends.

Data platforms

Monitoring the 57 HRSDG indicators requires at least 12 data systems to be functioning in each country (appendix pp 1–4). The most extensively used systems include nationally representative household surveys (NRHS) with a wide array of modules (including biomarker collection), civil registration and vital statistics (CRVS), and various administrative data systems. Alternatives exist for many preferred data systems, such as collecting verbal autopsy data through household survey or sample registration systems for CRVS. For some more complex indicators, such as deaths attributable to air pollution (indicator 3.9.1), multiple sources are required. For several indicators, preferred measurement methods are not direct measures from collected data and must be transformed or modelled for reporting on specific indicators. For example, for malaria incidence (indicator 3.3.3), due to the often low or variable disease notification rate in many high-burden settings, surveys that measure parasite presence in blood samples serve as the preferred measurement method. Overall, 13 HRSDG indicators should be based on complete CRVS, 20 on household surveys, nine on inputs from NRHS, 12 on administrative data, and three on data from multiple sources. CRVS, administrative records, and household surveys feature prominently as the preferred data systems for monitoring the HRSDGs (appendix p 5). CRVS, administrative systems, and disease registries can produce data on an annual basis, but time lags can affect the collation and reporting of data.

Data processing to address bias and enhance comparability

Meaningful comparisons over time and across or within countries require correcting for known data biases, as well as converting proxy or alternative measures to the reference case definition through a defined analytical approach (otherwise known as cross-walking). For example, ten indicators should use data from CRVS with medical certification of causes of death. But even among functional CRVS, recorded deaths must be corrected for completeness for vital registration system coverage, as well as highly variable certification and coding practices.²² The fraction of International Classification of Diseases (ICD) non-specific codes, or codes that cannot be an underlying cause of death (so-called garbage codes), ranges from 2·3% (in Singapore in 2016) to 83% (in the Maldives in 2005).²² Without correction, these biases can lead to overestimation or underestimation HRSDGs and then incorrect orderings of countries or subnational units.²²

For some HRSDGs, household surveys collect similar data, albeit in different ways. Alternative assays, items, or instruments must be cross-walked to the reference

measurement approach to ensure comparability across countries and over time. For malaria (indicator 3.3.3), data from prevalence surveys, or less often case notifications, can be based on thick and thin blood smears or rapid diagnostic tests, all of which feature different diagnostic sensitivities and specificities. Meaningful comparisons require adjusting for differences in test characteristics. Another example involves the duration of recall period used in surveys; for instance, many surveys on gender-based violence ask for lifetime recall, whereas others limit recall to a month or a year.^{23,24} Tuberculosis incidence (indicator 3.3.2) cannot be directly measured, so reported cases must be adjusted upwards by case-detection rates. Several approaches have been proposed to assess case-detection rates, including expert judgment, comparisons to prevalence surveys or triangulation using death rates, or prevalence in surveys.^{25–27} To ensure comparability, approaches used to triangulate data from multiple sources and across countries must be done consistently.

How data are processed can profoundly affect results, yet very little attention has been given to data processing and HRSDG measurement. For example, the IAEG-SDGs do not provide detailed guidance on data processing, and few scientific papers focus on data processing—a stark contrast to the substantial debates on data synthesis methods.^{28,29} In the absence of more explicit guidance, we would naturally expect heterogeneity in how national statistical offices, ministries of health, UN agencies, the GBD collaboration, and others have implemented data processing, undermining comparability across these efforts. Assessing bias in administrative data due to exclusion or misaligned incentives is an important aspect of data processing.³⁰

Data synthesis and imputation

Fundamental differences in opinion exist on how to move available processed data to final HRSDG values for a given country and year. Data synthesis and imputation should, in principle, account for both sampling and non-sampling error of measurements, synthesise multiple data points from different sources for the same indicator and year for a given country, and address the common problem of having no data for a given country and year. Although national authorities, UN agencies, the GBD collaboration, and various research groups use diverse approaches for data synthesis, these approaches can be roughly grouped into four families. In the following sections we discuss advantages and disadvantages of each approach.

Minimalist approach

With the minimalist approach, indicators are reported for years in which processed data are available. When more than one source for the same year is available, one source is selected on the basis of qualitative assessments of the quality of the different measurements. When different sources show large fluctuations over time (eg, stunting in South Africa),³¹ results are simply reported as they are.

The advantage of this approach is that producers of indicator results can point to a specific data source when questions emerge regarding the validity of results. However, there are two major disadvantages. First, when measurements are periodic—as all survey-based HRSDGs are—only a small subset of countries can be directly compared for a given year. Even for indicators that are based on CRVS or administrative systems, lags in data collation and reporting mean that results might only be available for 3–4 years before the year of data publication (eg, in 2019, the most recent data points could be for the years 2015–16). Second, sampling and non-sampling errors in measurements can lead to substantial, unexplained year-to-year fluctuations in indicator values. These fluctuations undermine the notion of accountability for change—a cornerstone of the HRSDG monitoring framework.

Minimalist approach plus the assumption of no change over time

Another approach starts with processed data and then increases the number of countries that can be compared by assuming indicator values are not changing over time since the most recent measurement.^{32–34} For example, a survey-based measure from 2011, if it is the most recent one, is used for 2018. Assuming no change in trend since the last available measurement is not credible for nearly all HRSDGs, as empirical evidence shows that most indicators have been improving globally. Assuming no change over time introduces considerable bias in reported comparisons; countries with more recent data will look better, all else being equal, than those with older measurements.

Expert data synthesis

The third approach tries to address the limitations of the minimalist approach by using a panel of experts to review available data and then determine the best estimate for each year. WHO/UNICEF Joint Reporting Process on vaccine coverage is an example of this approach,³⁵ wherein an expert group evaluates administrative data and survey data provided by countries and, on the basis of their assessment, annual vaccine coverage estimates are provided. An advantage of this approach is that it incorporates additional information that might not be captured by various quantitative sources (eg, vaccine stock-outs). Yet, disadvantages involve not being replicable by others—in fact, a different expert panel could easily come to diverging judgments. Lastly, the expert data synthesis approach does not follow the Guidelines on Accurate and Transparent Health Estimates Reporting (GATHER), which require well-documented statistical models for data synthesis to generate uncertainty estimates and allow replication.³⁶

Statistical data synthesis

The statistical data synthesis approach uses statistical models to synthesise processed data into a coherent

trend, applying the same model to estimate HRSDG indicator values for years with and without data, as well as generate uncertainty estimates for each year. The Inter-agency Group on Child Mortality Estimation is among the organisations applying this approach.³⁷ Advantages include enabling replicability by others, reporting of uncertainty, and producing estimates for a full time series, including years beyond the most recent processed data point. Some view this approach as potentially disadvantageous, such that using statistical models to generate estimates after an indicator's last measurement could misrepresent important trends or reversals due to data sparsity or lags in administrative data, or giving the false impression of having more data than what is available.³⁸ In addition, disagreements around how to best do statistical data synthesis for a given indicator are likely to occur and could lead to dissimilar models and results between analytical groups, even if all other analytic components—namely, data inputs and processing—are standardised. Finally, performance of statistical data synthesis generally improves by analysing data from all countries together; Bayesian statistical models often used in these efforts enhance predictive validity by borrowing strength over space and time.

Preferred approaches to data synthesis

Views about preferred approaches to data synthesis are as much philosophical as scientific. Some groups want to remain true to available observations and others would like to report indicator values that are as close as possible to what would have been observed with unbiased data systems and timely reporting. There is a substantial divide across and within organisations on the preferred approach: the World Bank largely follows the minimalist tradition for some indicators but statistical data synthesis for poverty estimation;³⁹ WHO uses the minimalist approach plus the assumption of no change over time for many tracer indicators in the universal health coverage service coverage measure (eg, prevalence of raised blood pressure),³⁴ expert synthesis for vaccine coverage,³⁵ and statistical data synthesis for malaria,⁴⁰ HIV,^{41,42} and tuberculosis.⁴³ GBD uses statistical data synthesis for all HRSDGs, with the exception of 17.19.2a (population census or registry status). There is a similar debate regarding estimates of uncertainty, particularly in terms of what types of uncertainty should be captured (ie, sampling and non-sampling errors), and for what results uncertainty should be reported. These are non-trivial areas for HRSDG measurement, especially since reporting of uncertainty estimates can identify instances in which more serious data gaps remain and inform investments in data collection and systems.

Country ownership

Using data to take action is much more likely when country governments, academia, and civil society fully understand and own every step of the measurement process, including data processing and data synthesis.

This notion of country ownership is also enshrined in UN Statistical Commission principles,⁴⁴ and is also why WHO has adopted the principle that Ministries of Health should be consulted on all indicator assessments. In an ideal world, every country would have high-quality, publicly available data reported with minimal lags from CRVS, periodic NRHS, and administrative health service data alongside other sources, such as National Health Accounts. Further, in this ideal setting, accepted international standards would exist to guide data processing and data synthesis, thereby enabling national statistical offices, ministries of health, academia, UN agencies, the GBD collaboration, and civil society groups to arrive at exactly the same answer for HRSDG monitoring. In this ideal case, the data would literally speak for themselves.

Unfortunately, our current reality is quite different. Due to myriad factors—sparse data, particularly in lower-income and middle-income countries, no international standards for data processing, and no accepted standards for data synthesis—different groups at both country and international levels can reasonably get quite different results on the HRSDGs. Discordant metrics from various actors committed to monitoring the HRSDGs is strong evidence of this mixture of problems. In some cases, national assessments of indicators differ from WHO or GBD assessments, with origins traceable to the data sources examined, data processing choices, and different approaches to data synthesis. To maintain comparability and foster accountability, data processing and data synthesis must be standardised. Data ownership must lie with data producers, which are often national governments. However, these data should also be made publicly available, as they are usually produced using public resources and are intended for the common good. Openly sharing data ultimately lays the foundation for the best health practices—those that are informed by the best possible science that draws from all available data and represents the best evidence base at a given point in time. This scenario is mutually beneficial to countries that share data (ie, they can put it to the best use) and to experts who aim to continuously improve the measurement of health outcomes and determinants. Promoting heterogeneity in data processing and data synthesis standards in the name of advancing country ownership is not helpful. Real country ownership will happen when national data systems are strengthened for data generation, analysis, and use, as well as when clear standards for each step of HRSDG monitoring are established. At that point, everyone will own the results since anyone examining the data will come to the same conclusion. Similar to other large-scale multi-country collaborative efforts, the GBD collaboration offers an example wherein transparent analysis and very broad participation encourages national ownership and use. The collaborative scientific model espoused by the GBD is one example of how to establish highly standardised approaches to data processing and synthesis while also fostering broad ownership.

Transparency

Expectations for transparency and replicability from the public, media, civil society, and scientific communities have been greatly enhanced in the past two decades. WHO led the creation of GATHER,³⁶ a much needed move towards transparency and a powerful way to improve trust across all segments in society in the validity of HRSDG monitoring. Although there has been great progress towards transparency, much more needs to happen. WHO estimates of child mortality, maternal mortality, and tuberculosis are fully compliant with GATHER, as are all outcomes produced by the GBD collaboration. However, many efforts that generate global health estimates are not compliant with GATHER. GATHER call for the public release of all input meta-data, but do not require publication of the actual values of input data. One key reason for this specification is that, often, national governments report data to WHO or other agencies but do not clarify if the data can be publicly shared. Some parts of WHO share data unless contributing governments give written prohibition, while other WHO departments do not share data unless permission is explicitly provided. WHO is moving towards a default opt-out position, such that data would be shared unless written agreements restrict sharing. This position would substantially enhance transparency and build trust in the global public health community. In some cases, governments restrict access and use of data in writing, posing challenges to transparency and replicability. In the future, transparency could be enhanced if the international community collectively agrees to only use data shared in the public domain for HRSDG monitoring.

Inequalities

The global SDG framework clearly emphasises the importance of addressing inequalities within countries across its goals, targets, and indicators.^{1,2,45} Monitoring HRSDG inequalities requires considerably more disaggregated data, which then necessitates the inclusion of appropriate equity stratifiers in NRHS or spatially disaggregated data from CRVS and administrative systems. However, examinations of HRSDG inequalities across indicators like socioeconomic status or minority groups (eg, race, ethnicity, and religion) have yet to widely occur, largely because these data are simply not available. An absence of practical and comparable approaches to assessing inequalities could mean that countries do not pay sufficient attention to reducing inequalities. The axis over which inequality varies the most (eg, sex, gender, age, sexual orientation, income, education, race, ethnicity, religion, and migrant status) will differ across countries. The global community is in a virtual impasse on inequality tracking for the SDGs: the choice appears to be to promote local inequality measurement using the most relevant axis of differentiation and lose all comparability and potentially accountability, or to inaccurately reflect the true nature of inequality in each country by using a standardised and

comparable measure along a single axis of differentiation. Inequality monitoring with disaggregated data is the first step towards reducing inequities (ie, inequalities deemed as unjust). The only imperfect option available today is to track geographical inequalities. UNICEF has focused on district-level inequalities in its EQUIST work,⁴⁶ and increasing number of HRSDGs have been analysed using model-based geostatistics.^{47–53} The most plausible prospect for adding an inequality dimension to HRSDG monitoring is to evaluate each indicator at the district or more disaggregated level, such as 5 km × 5 km pixels. Nonetheless, efforts to report on geographical inequalities must be clear that this is only one useful dimension of inequality measurement. Simultaneously, there needs to be investment in strengthening the capacity of data to enable reporting by other inequality dimensions.

Reporting

How indicators are reported generates substantial tension.⁵⁴ Accountability requires social engagement and political debate. In most societies, media (print, digital, and social) are essential components to accountability cycles.^{55–57} The media want to highlight rankings, poor performance, and, less often, success stories. By contrast, countries often prefer to not publicly discuss unfavourable comparisons. However, the most important thing is that HRSDG indicator reporting is clear and simple, and that the actual indicator values are readily available for countries to learn from others with comparably better performance. A transparent audit trail of every estimate that is published must be provided to build confidence in these numbers.

Recommendations

Our analysis of monitoring of the HRSDGs by national governments, WHO, the World Bank, the GBD collaboration, and others suggests some concrete steps forward to improve the timeliness, reliability, and validity of measurements. Pursuing these recommendations will hopefully lead to a situation in which the correlation of HRSDG indicator values will be much higher across different groups.

Strengthening national data collection capabilities

The international community, led by WHO, should actively seek to provide resources and technical assistance to countries to help them strengthen CRVS. Given the major role for various types of NRHS modules, including biomarkers in HRSDG monitoring, WHO should promote and streamline NRHS capability. WHO is actively considering launching a revised World Health Survey Plus data collection platform that would be tailored to the efficient measurement of the relevant HRSDG indicators. A regular, institutionalised, adequately resourced system of NRHS that includes biomarker collection would go a long way in also enabling better inequality analyses because these surveys would collect geolocated data for the

appropriate equity stratifiers. Better measurement and accountability must begin with improved primary data. Building national partnerships between the national statistical offices and the ministries of health will be key in this regard.

Data processing standards

Led by WHO, drawing on the experience of all the groups involved in HRSDG measurement, precise guidance should be developed for the processing of data that covers all key sources of bias, including assessing completeness of important registration, dealing with imprecise or impossible cause of death codes through redistribution algorithms, and cross-walking alternative case definitions, assays, instruments, and measurement methods to a defined reference approach. WHO has a clear leadership role in setting standards for health data; the ICD is a powerful example of this role. The normative role of WHO in setting data processing standards for HRSDG monitoring is extremely important.

Regularly updated best practice guidance on data synthesis

The international community should promulgate and adopt standard approaches to data synthesis reflecting current best practice. However, given that methods for data synthesis are rapidly evolving, principles by which improved methods are evaluated and adopted should also be clearly defined, such as enhanced out-of-sample predictive validity. National statistical offices and the myriad academic groups innovating in this area should be essential partners in this effort. Given widely divergent philosophical views on the right approach to data synthesis, these two extreme positions should be accommodated: the minimalist approach and the statistical data synthesis approach. Results can easily be produced using both approaches so that users can understand exactly what is coming from appropriately processed primary data and what is the result of the statistical data synthesis. This pragmatic compromise, producing both approaches, would then allow different groups to focus on their preferred set of results. As methods in this area evolve, international consensus can be built over a period of time.

Enhanced transparency guidelines

GATHER should be further strengthened. The scope of GATHER should cover all HRSDGs and inputs, including population numbers to the computation of the HRSDGs. For official HRSDG reporting, all data inputs, including data as they are used before and after processing, should be made publicly available. All code used in any analysis should be posted publicly. Strengthened GATHER, and their uniform application to HRSDG reporting championed by WHO, would have a ripple effect on other health analyses. If the norm for the community shifts to full data transparency and replicability, everyone will benefit. This scenario could become an effective antidote

to the unfortunate tendency towards the siloing of data that still exists in many settings.

Full global and national implementation of GATHER

Countries with analytical capacity should apply the standard for data processing and data synthesis to their own data in a fully transparent manner consistent with GATHER. National statistical offices or ministry of health statistics departments would do this entirely on their own or participate in global collaborations that fulfil GATHER, such as the GBD collaboration or interagency processes that estimate child and maternal mortality. Countries with insufficient analytical capacity need support to develop their analytical skills and infrastructure; WHO and the GBD collaboration with partners should make coherent efforts to advance this agenda. In the presence of sparse data, statistical data synthesis methods are more accurate when data from all countries are analysed jointly. A technical challenge remains in how to facilitate joint statistical data synthesis approaches while allowing groups in each country to experiment with and interrogate the methods.

Conclusion

The science of measuring the HRSDGs must be a guiding principle for sound measurement. Good measurement itself is not political—rather, the actions that are based on good measurement are political, as societies must make their own decisions and agendas, informed by the available data, national values, and social priorities. If the necessary data are collected, and the processing and synthesis steps standardised, results should not substantively vary on the basis of who conducts the analysis. These developments will naturally lead to a single, consistent, and transparently documented number for every HRSDG indicator. The scientific community has a crucial role in advancing the procedures and methods to support robust measurement of the HRSDGs.

Contributors

SA, SC, RL, and CJLM prepared the first draft and finalised the manuscript on the basis of comments and reviewer feedback. SS, NY, MR, MFM, EV, AM, and LD provided critical commentary on, and content for, the manuscript over the stages of manuscript finalisation.

Declaration of interests

We declare no competing interests.

Acknowledgments

The authors thank Nancy Fullman and J Everett Mumford for their contributions to this Viewpoint.

References

- 1 UN. Transforming our world: the 2030 Agenda for Sustainable Development. 2015. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication> (accessed July 7, 2019).
- 2 UN. Resolution adopted by the General Assembly on work of the statistical commission pertaining to the 2030 Agenda for Sustainable Development (A/RES/71/313). 2017. <https://unstats.un.org/sdgs/indicators/indicators-list/> (accessed July 7, 2019).
- 3 Nilsson M, Griggs D, Visbeck M. Policy: map the interactions between Sustainable Development Goals. *Nature* 2016; 534: 320–22.
- 4 International Council for Sciences (ICSU), International Social Science Council (ISSC). Review of targets for the Sustainable Development Goals: the science perspective. 2015. <https://council.science/cms/2017/05/SDG-Report.pdf> (accessed July 7, 2019).
- 5 WHO. World Health Statistics 2016: monitoring health for the SDGs. Geneva: World Health Organization, 2016.
- 6 Lim SS, Allen K, Bhutta ZA, et al. Measuring the health-related Sustainable Development Goals in 188 countries: a baseline analysis from the Global Burden of Disease Study 2015. *Lancet* 2016; 388: 1813–50.
- 7 UN Sustainable Development Knowledge Platform. SDG reviews from member states of the United Nations and specialized agencies. <https://sustainabledevelopment.un.org/memberstates> (accessed July 7, 2019).
- 8 UN Sustainable Development Knowledge Platform. Voluntary national reviews database. <https://sustainabledevelopment.un.org/vnrs/> (accessed July 7, 2019).
- 9 WHO. World Health Statistics 2017: monitoring health for the SDGs. Geneva: World Health Organisation, 2017.
- 10 WHO. World Health Statistics 2018: monitoring health for the SDGs. Geneva: World Health Organisation, 2018.
- 11 WHO. World Health Statistics 2019: monitoring health for the SDGs. Geneva: World Health Organization, 2019.
- 12 World Bank. Atlas of the Sustainable Development Goals 2017: from the World Development Indicators. 2017. <https://openknowledge.worldbank.org/handle/10986/26306> (accessed July 7, 2019).
- 13 World Bank. Atlas of Sustainable Development Goals 2018: from World Development Indicators. 2018. <https://openknowledge.worldbank.org/handle/10986/29788> (accessed July 11, 2019).
- 14 Fullman N, Barber RM, Abajobir AA, et al. Measuring progress and projecting attainment on the basis of past trends of the health-related Sustainable Development Goals in 188 countries: an analysis from the Global Burden of Disease Study 2016. *Lancet* 2017; 390: 1423–59.
- 15 Lozano R, Fullman N, Abate D, et al. Measuring progress from 1990 to 2017 and projecting attainment to 2030 of the health-related Sustainable Development Goals for 195 countries and territories: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; 392: 2091–138.
- 16 Sustainable Development Solutions Network (SDSN). SDG index and dashboards—global report. New York, NY: SDSN, 2016.
- 17 Sachs J, Schmidt-Traub G, Kroll C, Durand-Delacré D, Teksoz K. SDG index and dashboards report 2017. New York, NY: Bertelsmann stiftung and sustainable development solutions network, 2017.
- 18 Bertelsmann Stiftung, Sustainable Development Solutions Network. SDG index and dashboards report 2018. Global responsibilities: implementing the goals. 2018. <http://www.sdgindex.org/assets/files/2018/01%20SDGS%20GLOBAL%20EDITION%20WEB%20V8%20060718.pdf>. (accessed May 7, 2019).
- 19 Measuring progress towards the Sustainable Development Goals—SDG tracker. Our world data. <https://sdg-tracker.org/> (accessed July 5, 2019).
- 20 WHO. Toward a global action plan for healthy lives and well-being for all: uniting to accelerate progress towards the health-related SDGs. Geneva: World Health Organization, 2018.
- 21 WHO. Thirteenth general programme of work 2019–2023: promote health, keep the world safe, serve the vulnerable. Geneva: World Health Organization, 2018.
- 22 Roth GA, Abate D, Abate KH, et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* 2018; 392: 1736–88.
- 23 Ellsberg M, Heise L. Researching violence against women: a practical guide for researchers and activists. Geneva: World Health Organization, 2005.
- 24 Schwartz MD. Methodological issues in the use of survey data for measuring and characterizing violence against women. *Violence Against Women* 2000; 6: 815–38.
- 25 Dye C, Watt CJ, Bleed DM, Williams BG. What is the limit to case detection under the DOTS strategy for tuberculosis control? *Tuberculosis* 2003; 83: 35–43.

- 26 Borgdorff MW. New measurable indicator for tuberculosis case detection. *Emerg Infect Dis* 2004; **10**: 1523–28.
- 27 Claassens MM, van Schalkwyk C, Dunbar R, Ayles H, Beyers N. Patient diagnostic rate as indicator of tuberculosis case detection, South Africa. *Emerg Infect Dis* 2016; **22**: 535–37.
- 28 Boerma T, Victora C, Abouzahr C. Monitoring country progress and achievements by making global predictions: is the tail wagging the dog? *Lancet* 2018; **392**: 607–09.
- 29 AbouZahr C, Boerma T, Hogan D. Global estimates of country health indicators: useful, unnecessary, inevitable? *Glob Health Action* 2017; **10**: 1290370.
- 30 Friebel R, Silverman R, Glassman A, Chalkidou K. On results reporting and evidentiary standards: spotlight on the Global Fund. *Lancet* 2019; **393**: 2006–08.
- 31 Akombi BJ, Agho KE, Hall JJ, Wali N, Renzaho AMN, Merom D. Stunting, wasting and underweight in Sub-Saharan Africa: a systematic review. *Int J Environ Res Public Health* 2017; **14**: e863.
- 32 Boerma T, Requejo J, Victora CG, et al. Countdown to 2030: tracking progress towards universal coverage for reproductive, maternal, newborn, and child health. *Lancet* 2018; **391**: 1538–48.
- 33 Hogan DR, Stevens GA, Hosseinpoor AR, Boerma T. Monitoring universal health coverage within the Sustainable Development Goals: development and baseline data for an index of essential health services. *Lancet Glob Health* 2018; **6**: e152–68.
- 34 WHO. Primary health care on the road to universal health coverage: 2019 monitoring report. Geneva: World Health Organization, 2019.
- 35 WHO. WHO/UNICEF joint reporting process. WHO. https://www.who.int/immunization/monitoring_surveillance/routine/reporting/reporting/en/ (accessed July 5, 2019).
- 36 Stevens GA, Alkema L, Black RE, et al. Guidelines for Accurate and Transparent Health Estimates Reporting: the GATHER statement. *Lancet* 2016; **388**: e19–23.
- 37 WHO. MCEE-WHO methods and data sources for child causes of death, 2000–2017. Geneva: World Health Organization, 2018.
- 38 Tichenor M, Sridhar D. Metric partnerships: global burden of disease estimates within the World Bank, the World Health Organisation and the Institute for Health Metrics and Evaluation. *Wellcome Open Res* 2019; **4**: 35.
- 39 World Bank. Poverty and shared prosperity 2018: piecing together the poverty puzzle. Washington, DC: World Bank, 2018.
- 40 WHO. World malaria report 2018. Geneva: World Health Organization, 2018.
- 41 UNAIDS. Global AIDS update 2018. Miles to go: closing gaps, breaking barriers, righting injustices. Geneva, Switzerland: UNAIDS, 2018.
- 42 UNAIDS. UNAIDS data 2018. Geneva, Switzerland: UNAIDS, 2018.
- 43 WHO. Global Tuberculosis Report 2018. Geneva: World Health Organization, 2018.
- 44 UN. Resolution adopted by the General Assembly on Fundamental Principles of Official Statistics (A/RES/68/261). New York, NY: United Nations, 2014.
- 45 Sustainable Development Solutions Network (SDSN). Indicators and a monitoring framework for the sustainable development goals: launching a data revolution. New York, NY: SDSN, 2015.
- 46 UNICEF. EQUIST: equitable strategies to save lives. <https://www.equist.info/> (accessed June 13, 2019).
- 47 Golding N, Burstein R, Longbottom J, et al. Mapping under-5 and neonatal mortality in Africa, 2000–15: a baseline analysis for the Sustainable Development Goals. *Lancet* 2017; **390**: 2171–82.
- 48 Osgood-Zimmerman A, Milliar AI, Stubbs RW, et al. Mapping child growth failure in Africa between 2000 and 2015. *Nature* 2018; **555**: 41–47.
- 49 Mosser JF, Gagne-Maynard W, Rao PC, et al. Mapping diphtheria-pertussis-tetanus vaccine coverage in Africa, 2000–16: a spatial and temporal modelling study. *Lancet* 2019; **393**: 1843–55.
- 50 Weiss DJ, Lucas TCD, Nguyen M, et al. Mapping the global prevalence, incidence, and mortality of *Plasmodium falciparum*, 2000–17: a spatial and temporal modelling study. *Lancet* 2019; **394**: 322–31.
- 51 Battle KE, Lucas TCD, Nguyen M, et al. Mapping the global endemicity and clinical burden of *Plasmodium vivax*, 2000–17: a spatial and temporal modelling study. *Lancet* 2019; **394**: 332–43.
- 52 Dwyer-Lindgren L, Cork MA, Sligar A, et al. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. *Nature* 2019; **570**: 189–93.
- 53 Burstein R, Henry NJ, Collison ML, et al. Mapping 123 million neonatal, infant and child deaths between 2000 and 2017. *Nature* 2019; **574**: 353–8.
- 54 Murray CJL. Choosing indicators for the health-related SDG targets. *Lancet* 2015; **386**: 1314–17.
- 55 Bowles N, Hamilton JT. Transparency in politics and the media: accountability and open government. Oxford: I B Tauris, 2013.
- 56 Camaj L. The media's role in fighting corruption: media effects on governmental accountability. *Int J Press* 2013; **18**: 21–42.
- 57 Olukoyun A. Media accountability and democracy in Nigeria, 1999–2003. *Afr Stud Rev* 2004; **47**: 69–90.

© 2019 Elsevier Ltd. All rights reserved.