Report of the online survey to inform GATHER

To inform the development of the Guidelines for Accurate and Transparent Health Estimates reporting (GATHER), an online survey was carried out to obtain feedback on a list of potential reporting items. The list of potential reporting items was developed during 2013 and 2014, starting with a draft discussion paper circulated at a Reference Group for Health Statistics meeting at the World Health Organization (WHO), Geneva, in December 2013. All reporting items suggested at the December 2013 Reference Group meeting, items suggested by working group members in a series of e-mail and teleconference exchanges in the fall of 2014, and all relevant reporting items from the PRISMA guidelines for systematic reviews were included in this comprehensive set of 33 potential reporting items. These items were organized into five groups: context and design, data inputs, analytic methods, results, and discussion.

The online survey was designed to obtain feedback from a broader community of researchers and users of health estimates on the comprehensive set of potential reporting items. The online survey was developed using LimeSurvey. The complete text of the survey (i.e. screenshots) is included as an annex to this report. In summary, the survey first introduced the GATHER working group and project, the aim of the survey, the scope of the guidelines, and gave some advice on how to evaluate reporting items. The respondent’s country of residence and experience producing global health estimates were queried. Finally, the respondent was asked to evaluate each potential reporting item. For each reporting item, the following question was asked: “Should this item be in the reporting checklist?” with four response options: require, recommend, exclude, no opinion. An open text comment field was included for each item and at the end of the survey.

The survey was active from 20 January to 12 February 2015. Our strategy was to widely disseminate the survey in order to obtain feedback from a broad group of stakeholders. A key group of 57 academics who have published global health estimates (and who were not members of the working group) were identified and invited to respond on January 20th, with a reminder on January 27th. In addition, all working group members shared the survey in their respective networks (which likely resulted in cross-posting). Recipients contacted in this way included Institute for Health Metrics and Evaluation staff involved in global burden of disease estimation, United Nations expert groups such as the Inter-agency Group for Child Mortality Estimation (IGME), Maternal Mortality Estimation Inter-agency Group (MMEIG) and the UNAIDS Reference Group on Estimates, Modelling and Projections, and staff at the WHO, United Nations Children’s Fund (UNICEF), International Agency for Research on Cancer (IARC), and the United Nations Population Division who prepare or use global health estimates.

We did not collect personal identifying information on the respondents. We received 118 responses, of which 79 (67%) were complete. 110 (93%) responses were from residents of high-income countries, with the most responses residents of USA, UK and Switzerland. 88 (75%) of the respondents reported they had led or participated in a study that calculated global health estimates. For most questions and overall, there was no significant difference between the evaluations of individuals who had and who had not previously participated in such a study.

---

3 www.limesurvey.org
At the consensus meeting held in London, United Kingdom, in February 2015, a histogram showing the multiple-choice responses and a summary of the free text responses (both results follow) where shown on screen to guide the discussion. In addition, each working group member received the complete record of the free text responses.

The reporting items included in the survey were revised in light of the survey results and the discussions at the meeting. In addition, the working group decided to revise the scope of studies included in GATHER (to exclude studies of intervention coverage) after the online survey was completed.
# Context and design

## 1 Abstract

Summarize the data inputs, including inclusion/exclusion criteria, the analytic methods, the results, and the limitations of the study in the abstract (if an abstract is provided).

**Explanation** An abstract provides a general overview of the study that should be informative to both technical and lay readers.

**Survey response:**

![Frequency chart](chart.png)

**Comments from online survey:**

- Many comments were made that word limit restrictions would not allow for complete reporting in the abstract. For example, one person commented “the abstract is not the best place to report on inclusion/exclusion criteria, data inputs and so on, because the need to keep it short will result in very general and probably useless information”
- Other points suggested for inclusion in the abstract were: objectives/aim, conclusions, uncertainty measures, limitations, results of all primary objectives, and validations carried out.

## 2 Purpose

Describe the purpose of the estimates (e.g., priority setting, program evaluation) in the context of what is already known.

**Explanation** The purpose of a set of health estimates may depend on current debates in public health practice, and this purpose may inform analytic choices. The potential uses of a set of estimates may also depend on the amount of underlying data that are available to make the estimates. The analysts should state why the estimates were made and what uses are appropriate given the data and methods that were used.
3 Aim Provide an explicit definition of the health indicator that is described and the populations for which estimates are made.

Explanation: Include, as relevant, diagnostic criteria, target demographic group, and geographic areas for which estimates were made.

Survey response:

Comments from online survey:

• One person suggested merging with ‘purpose’.
• One person mentioned that data used might not match the indicator/populations in the stated aim.
Data inputs

4 Data search

**Explanation:** In the medical field, a systematic data search usually includes a review of the published scientific literature. However, health data are often found in the grey literature, including health surveys, government reports, and databases maintained by international organizations and academic groups. Researchers should specify which databases were searched. They should also mention whether other search strategies were used, such as contacting professional networks for unpublished data. This allows readers to assess the comprehensiveness and completeness of the search. This reporting item refers to measures of the health indicator being estimated.

Survey response:

![Bar chart showing frequency of responses to data search question]

Comments from online survey:

- Several commenters weren’t sure that this item made sense for this kind of data collection, e.g. if a database or gold-standard national survey is used.
- Two comments suggested following the PRISMA guideline on systematic reviews for this item.
- One person did not understand this item – the commenter did not expect that any data would be extracted from reports/articles.

5 Inclusion/exclusion criteria

**Explanation:** Consistent application of defined inclusion/exclusion criteria is a key characteristic of a systematic review. Information on these criteria will allow the reader to understand the quality of data used, as well as potentially informative data that were excluded. The authors should specify publication / data years included, languages, study designs, and demographic groups included, as relevant. This reporting item refers to measures of the health indicator being estimated.
Survey response:

Comments from online survey:

- Two people felt this would not be possible
- One person thought it did not make sense if an existing database is used
- One person thought language constraints should be specified, and one felt these should be pre-specified.

6 Ad-hoc exclusions

If any data sources which met inclusion criteria were eventually excluded, identify the data sources and the reason(s) for exclusion.

Explanation: Data sources which report implausible values are sometimes excluded on an ad-hoc basis. Ad-hoc exclusions should be explicitly identified so that the study can be replicated. This reporting item refers to measures of the health indicator being estimated.

Survey response:

Comments from online survey:

- Comments were generally supportive. One person suggested stating if there are no ad-hoc exclusions.

7 Data extraction

List and define all variables for which data were sought.

Explanation: This reporting item refers to characteristics of each data source, such as representativeness, data collection method, or demographic characteristics of the study population.
Survey response:

Comments from online survey:

- One person thought this was too cumbersome, if a wide range of variables are sought from a set of studies.
- Several people mentioned that points 7, 8 and 9 are not relevant if an online database is used.

8 Data extraction
Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate).

Explanation: Studies that involve a substantial effort to extract data from a variety of sources, thus creating a new data base, should provide information about how the data were extracted from existing sources.

Survey response:

9 Data extraction
Specify whether authors were contacted.

Explanation: When reviewing potential data sources to inform a global health estimate, researchers may identify existing studies that collected useful data, but did not publish those data in a relevant form. In these cases, researchers should consider contacting the authors of the study to request these data so as to improve the accuracy of the estimates.
Survey response:

Comments from the online survey:

- Several mention that the specific data sources for which authors are contacted should be identified.
- One person thought that if any authors are contacted, all should be contacted.

10 Data description

Identify and describe any types of input data that provide a potentially biased estimate of the quantity of interest.

Explanation: Input data may provide a biased estimate of the health indicator of interest. For example, case definitions or diagnostic criteria may vary between data sources. Readers may want to know whether estimates are informed by potentially biased input data. This reporting item refers to measures of the health indicator being estimated.

Survey response:

Comments from the survey:

- Several commenters thought the discussion of potential sources of bias was confusing.

11 Data description

Provide a complete list of data sources used and their key characteristics. Key characteristics may include, as relevant, Key characteristics potentially include: 1) reference information; 2) sampling strategy; 3) population represented (e.g. national, subnational, occupation-based; 3) sex and age range; 4) year of data collection; 5) diagnostic criteria or measurement method; 6) sample size; 7) assessor’s qualifications; and 8) potential sources of bias.
**Explanation:** This is vital information for understanding the basis of the estimates. For example, an estimated trend may be produced on the basis of a few studies of variable quality in a particular country. By defining the data used and their key characteristics, the user can understand the basis of the estimate for each population. This reporting item refers to measures of the health indicator being estimated.

**Survey response:**
Other sources of bias

Comments from online survey on data description (items 10 and 11):

- Other possible descriptors suggested include: case detection method (active vs. passive), point vs. period prevalence, and if period, length of period; standard error and/or confidence interval, response rate,
- Several people thought ‘potential sources of bias’ was confusing or unhelpful wording.

12 Data description

Provide a graphical representation of data availability in space and/or time.

Explanation: In some cases, the list of data sources is lengthy. A graphical representation of data availability can illuminate whether there are geographical areas or time periods where data are particularly sparse.

Survey response:

NB. Responses from analysts and non-analysts were significantly different. Non-analysts were more likely to exclude this item, and analysts were more likely to recommend it.

Comments from online survey:

- There were several comments in support of this item. Several thought it would be too hard to implement and a list of data sources are sufficient.

13 Covariates and other data inputs

Describe and give the source of other data inputs, such as covariates, used to generate health estimates.
Explanation: Often data that are not measures of the indicator being estimated are used to make estimates, such as covariates or model parameters in a natural history model. The sources for these data should be fully described.

Survey response:

Comments from online survey on covariates:

- Two people commented in support of this item.

14 Covariates and other data inputs

For each data input, identify whether it met the relevant reporting guidelines (e.g., PRISMA guidelines for meta-analyses, these guidelines for health estimates, CONSORT guidelines for trial data).

Explanation: Covariates or model parameters used in these studies may not adhere to the relevant reporting guidelines. These parameters may have an important influence on the health estimates.

Survey response:

Comments from online survey on covariates:

- Several people mentioned that adherence to guidelines does not ensure quality. One stated that it may not be possible to tell whether the data adhere to the guidelines.
- Many people mentioned that this would be extremely burdensome.
15 Replicability  Make all data inputs available. If confidential data were used, all remaining data should be made available and results of analyses using the subset of reported data should be given. Include public access to all input data, all covariates/model parameters. If any input data are confidential, study results should be provided using only the public-access input data.

Explanation: Researchers need access to data sets to promote replication and improvement in the methods. When some data are restricted access, the comparison of the full and available data sets let other analysts have a better understanding of the importance of the inclusion of the restricted data sets.

Survey response:

* Excluding confidential data. Responses to this question were significantly different between data analysts and non-analysts. Analysts were more likely to require this item, and non-analysts were more likely to exclude.

Comments from online survey:

- Many people were supportive of open access to input data. Several suggested a link to data providers in case of confidential input data.
- One person mentioned that it would be a lot of work for GBD to rerun without the confidential data.
- Many mentioned legal issues, i.e., a data provider does not allow a user to republish the statistic for further use by others. One person stated this should not be an impediment to research.
- One person mentioned that data users should provide contact information.
- Several people mentioned that data access/replicability may be hard to organize if the dataset are large and models are complex.
- One person mentioned new guidelines from donors e.g. BMGF, Wellcome Trust, and from PLoS on public access to data.
Analytic methods

16 Analytic overview

Provide a lay description of the analytic methods, including what assumptions are implied by the methods.

Explanation: All users (technical and non-technical) should be able to understand the analysis. In particular, there may be assumptions embedded in models about how estimates are made, i.e., linear time trends or time trends based on covariates only. Some assumptions may particularly affect out-of-sample estimates. These should be clear to the user.

Survey response:

![Bar chart showing survey responses]

Comments from online survey:

- Three people felt ‘lay’ should be dropped, while two others thought lay/comprehensible description would be critical. One of the people who objected to the term ‘lay’ said that it feels like a requirement to ‘dumb down’ a manuscript.

17 Analytic overview

If the methods involves multiple analytic steps, provide a diagram showing model structure and/or analytic steps.

Explanation: Studies often involve multiple analytic steps. These steps may be difficult to describe in text, while a flow diagram would be more understandable.

Survey response:

![Bar chart showing survey responses]
Comments from online survey:

- Two people commented that a diagram is not necessarily the most helpful way to describe methods.

18 Data Preparation

**Description:** Describe any pre-processing of input data.

**Explanation:** Examples of data pre-processing include methods for calculation of population-level indicators from data on individuals, or pre-processing of covariate data. These steps should be described to allow for replication of the study.

Survey response:

![Bar chart]

Comments from online survey:

- One person thought the description should be simple. A second person wanted complete documentation, including code.
- It will be important to clarify how this item differs from item 23.

19 Model Selection

**Description:** Describe the models that were considered during the model selection process and how the final model(s) were chosen. This includes providing a list of covariates evaluated, if covariates were used.

**Explanation:** This will give an indication of the quality of the model-building process.

Survey response:

![Bar chart]

Comments from the online survey:

- One person felt this was too cumbersome as one goes through many models while debugging.
• On this item and item 20, one person mentioned that these are unclear since ‘model’ has many meanings.

20 Model selection
Report how model performance was evaluated.

Explanation: Report whether and how model assumptions were checked. Report quantitative measures of goodness-of-fit and of cross-validation (if carried out). If k-fold cross validation is carried out, report the types of training and test sets used and the quantitative measures of goodness-of-fit used for evaluating model performance. Report if estimates were verified by comparisons between different models. These reports will give an indication of the thoroughness of the model-building process.

Survey response:

21 Model selection
Provide a graphical representation of model fit, i.e., plots comparing model output to adjusted and/or unadjusted data inputs.

Explanation: Many users will want to understand how well the model fits the available data; a graphical representation is accessible to all users regardless of their statistical background. This can be included in an appendix.

Survey response:

Comments from online survey:

• One person said that graphs are not the correct way to assess fit. The explanation can be rephrased to say that such plots show how the modelled estimates compare to the input data.
22 Methods description

Provide a technical description of the model(s) finally used to calculate health estimates.

Explanation: The technical description should be sufficient for replication. It should include mathematical formulae or statistical code for each analytical step and specification of all model parameters. The statistical package(s) used for the analysis should be named.

Survey response:

Comments from online survey:

- Several people thought the technical model description should go in an annex. Several people commented that the model code should also be made available.

23 Methods description

For each potentially biased input data type, report whether methods were developed to adjust the data or describe the reason for failing to adjust the data.

Explanation: Adapted from PRISMA reporting guidelines for meta-analyses.

Survey response:

Comments from online survey:

- Some commenters found the term ‘potentially biased input data type’ confusing.
24 Sensitivity analyses

Describe methods of sensitivity analyses (i.e. additional analyses) and findings, if done.

Explanation: All estimation processes involve assumptions, including inclusion criteria for data and the functional form of a model. The degree to which the final values of the estimates depend on these assumptions should be explored and reported.

Survey response:

![Bar chart](image)

Comments from online survey:

- Three people thought sensitivity analyses should not be required.

25 Uncertainty analysis

Specify which sources of uncertainty are reflected in the uncertainty range.

Explanation: Often uncertainty ranges reflect only a subset of sources of uncertainty. The sources of uncertainty which are reflected in the uncertainty range should be listed (e.g., sampling errors in input data, non-sampling errors in input data, uncertainty from multiple analytic steps, and model uncertainty may be reflected in final uncertainty ranges).

Survey response:

![Bar chart](image)

Comments from online survey:

- Several commenters were supportive. One person added that sources of uncertainty which are not reflected should also be listed.
26 Replicability Make all statistical code available.

Explanation: Public access to code ensures that results are replicable. Reproducing an analysis from a description of methods in a paper is generally difficult and often impossible because there are too many details to be comprehensively described in the methods. For most researchers providing a general description of the approach with key assumptions, along with access to the code for more specific details would be easier to do.

Survey response:

![Survey Graph]

Comments from online survey:

- Several people expressed concern about the need retain code to publish and copyrights. Most of these people agreed that in theory sharing is a good idea.
- One person commented that the calculations may be carried out in a spreadsheet. One person thought sharing code was not necessary if a published package/command is used. One person clarified that it may not be necessary to share all ‘data-cleaning’ code, just the input dataset and the model itself.
- Several people were very supportive of this item in their comments.
Results

27 Results Provide health estimates in machine-readable format, disaggregated at the level used for reporting.

Explanation: This is important for straightforward use of estimates in other applications. This would not be necessary in a situation where only a few estimates are made, that is, for 5-10 world regions and one time period.

Survey response:

Comments from online survey:

- One person was concerned that the ‘machine-readable’ requirement was onerous.
- Other comments were supportive.

28 Uncertainty analysis Provide and interpret uncertainty ranges (e.g., 95% or 90% uncertainty intervals).

Explanation: Uncertainty ranges should reflect the largest source(s) of uncertainty.

Survey response:

Comments from online survey:

- There was confusion about the term ‘interpret’ – many thought it referred to political decision-making.
- Several people felt that UI’s should be recommended, but that they should not be required, as they did not want this requirement to prevent research from being published.
Several people agreed that UI’s should be required and described. One person thought that guideline should not make any statement about including the largest source of uncertainty.

29 Evidence basis of results

Categorize estimates for each population according to the amount and/or quality of input data from the same population that informed the estimates.

Explanation: Most comprehensive global health estimates include estimates which are informed by variable amounts/quality of underlying input data. For example, there may be no data available for population A, one low-quality data source available for population B, and several high-quality data sources available for population C. Out-of-sample predictions (such as the one for population A) may be used to identify likely health priority areas in populations for which measured health data are not available, and can provide insight into probable trends in health statistics where trend data are not available or are inconsistent. However, it is important for users to recognize the limitations of these estimates. For example, out-of-sample health statistics are not a substitute for measurements when evaluating disease programs or health systems reforms. If possible, health estimates with missing or low-quality input data should be identified in graphical or tabular presentations of health estimates, e.g. by using a different color or font to identify these statistics.

Survey response:

Comments from online survey:

- One person suggested ‘recommend’ since the input data should be reported.
- Several people thought this would be a useful recommendation, and evidence availability by country income/region should also be reported.

30 Time trends

If data support estimates of trends over time, provide estimates for multiple years (vs. a single point in time).

Explanation: Some groups make estimates for a specific year on a routine basis. Then others (including the press, general public and even scientists) have combined these multiple point estimates to produce trends which in fact were incorrect as both data and methods had changed between the different rounds of estimation. If all sets of estimates include trends over time, it would clarify to the user that different sets of estimates should not be compared.
Survey response:

![Bar chart showing survey responses: require, recommend, no opinion, exclude, no response.]

Comments from online survey:

- Several people felt this item was about conduct, not reporting. One person mentioned that this would have a cost.
- Others commented that trends are useful.
Discussion

31 Comparison with prior estimates

If updating a previous set of estimates, the analyst should report the results of a decomposition analysis that explains changes.

Explanation: Updated estimates may change due to updated datasets, methods changes, or both. The analyst should perform an analysis crossing datasets and methods to identify why estimates have changed. This will allow users to understand the cause of any differences between the two rounds of estimates.

Survey response:

Comments from online survey:

- Several commenters were supportive of this item.
- Several people thought that a complete decomposition was too time-intensive and not necessarily worth the effort.

32 Limitations

Describe limitations in data inputs and analytic methods.

Explanation: Important assumptions should be discussed, including the effects of analyst choices regarding data inclusion, selection of a model (or models) to generate health estimates, and methods for potentially biased data. Specifically identify major sources of uncertainty, including any that were not captured by the uncertainty range.

Survey response:
Comments from online survey:

- Two people pointed out that authors may not think of all limitations.
- Some suggested a particular focus on how the limitations translate to quality of estimates and their suitability for policy decisions.

33 Research needs

Describe studies that should be carried out to improve future estimates.

Explanation

Authors of a new study are in a unique position to identify potential improvements to data and methods that would improve future estimates.

Survey response:

![Bar chart showing frequency of responses]

Comments from online survey:

- Some felt that papers on health estimates aren’t the right ones to make these comments.
- One person commented that the EPICOT framework could be useful to make this item more specific.
Annex 1: Complete screenshots of the online survey
Developing guidelines for reporting global health estimates: the GATHER project

The Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER) working group is seeking feedback on how global health estimates should be reported. In this survey, you will be asked your opinion on what information should accompany new publications of global health estimates in journal articles or in reports. Your responses will be used to develop the GATHER reporting guidelines for global health estimates. Read on for more details.

How should analyses of global health estimates be reported?

Global, regional, and country estimates of health indicators are used for monitoring health and guiding resource allocation. Because of major data gaps and measurement challenges, models are often used to fill data gaps. Users of these estimates need basic information about the data and the methods used in order to interpret these estimates. To address this need, the World Health Organization convened a working group, composed of researchers and journal editors who have published global health estimates, to define best practices for documenting studies that report global health estimates (members listed below). This working group is developing the Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER).

What is the aim of this survey?

The GATHER working group has compiled a list of potential items to include in a reporting checklist. We are seeking feedback from users and researchers on which items should be reported with every new set of health estimates that is published. The reporting items are organized into five groups:

- context and design,
- data inputs,
- analytic methods,
- results, and
- discussion.

In the next screen, you will answer some basic questions about your experience with global health estimates. You will then see one screen for each of the five groups of items listed above, in which you will be asked to rate the importance of each potential reporting item.

What are global health estimates?

Global health estimates include all 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 and injuries; and indicators of health determinants, such as exposure to risk factors or intervention coverage.

GATHER aims to define best practices for studies that synthesize evidence from multiple sources to quantitatively describe past and current population health. These studies often include a comparison among multiple populations, over time or by place of residence. They include studies that disaggregate disease and injuries by underlying cause as defined by the International Classification of Disease [ICD] rules and those that attribute disease and injury to their determinants, such as calculations of the number of deaths attributable to tobacco smoking. These reporting guidelines are not meant to cover reports of a health indicator from a single study or data source, such as a health examination survey.

Health determinants can range from proximal determinants of health, such as behaviors like tobacco smoking that have a direct effect on incidence of disease and mortality, to intermediate determinants of health, such as availability of essential
medicines, to distal determinants of population health, such as wealth inequality in a population. Of the universe of health determinants, these reporting guidelines are being developed for estimates of health service coverage and risk factors that generally fall under the purview of the health system.

**How should reporting items be evaluated?**

Reporting should be sufficient for a knowledgeable reader to be able to evaluate the quality of a global health estimate. Reporting of estimates should serve the needs of their two primary audiences, users and researchers. Users will need information about data sources used and analytical methods, including understanding limitations of estimates, in a way that can be interpreted without advanced training in statistics. Researchers will require a higher degree of detail about methods, so that they can reproduce studies and advance methods.

Note that the aim of these reporting items is to improve reporting of global health estimates, not to provide guidance on how to calculate global health estimates. Also note that some of the reporting required by the items in this survey are lengthy and would need to be included in an online supplement.

**GATHER working group members:**

Leontine Alkema, National University of Singapore
Ties Boerma, World Health Organization
Robert Black, Johns Hopkins Bloomberg School of Public Health
Majid Ezzati, Imperial College London
John Grove, Bill & Melinda Gates Foundation
Daniel Hogan, World Health Organization
Richard Horton, The Lancet
Joy Lawn, London School of Hygiene and Tropical Medicine
Ana Marusic, University of Split
Christopher Murray, Institute for Health Metrics and Evaluation, University of Washington
Igor Rudan, University of Edinburgh
Joshua Salomon, Harvard School of Public Health
Gretchen Stevens, World Health Organization
Theo Vos, Institute for Health Metrics and Evaluation, University of Washington
Neff Walker, Johns Hopkins Bloomberg School of Public Health
Developing guidelines for reporting global health estimates: the GATHER project

Personal information

* Have you led or participated in a study that calculated global health estimates?
  - Yes
  - No

* Please select your country of residence.
  Choose one of the following answers
  - Please choose...
Developing guidelines for reporting global health estimates: the GATHER project

Context and design

Should each reporting item be a required checklist item, recommended checklist item, or should the reporting item be excluded from the checklist?

Tip: to navigate to another section of this survey, use the Question Index on the right of the screen or the previous/next button at the bottom of the page (you may need to scroll down).

Abstract

Summarize the data inputs, including inclusion/exclusion criteria, the analytic methods, the results, the conclusions, and the limitations of the study in the abstract (if an abstract is provided).

<table>
<thead>
<tr>
<th>Should this item be in the reporting checklist?</th>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
</table>

Adapted from PRISMA reporting guidelines for meta-analyses.

Any additional comments on reporting in an abstract:

Purpose

Describe the purpose of the estimates (e.g., priority setting, program evaluation) in the context of what is already known.

<table>
<thead>
<tr>
<th>Should this item be in the</th>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
</table>

https://extranet.who.int/dataform/survey/index
The purpose of a set of health estimates may depend on current debates in public health practice, and the purpose may inform analytic choices. The potential uses of a set of estimates may also depend on the amount of underlying data that are available to make the estimates. The analysts should state why the estimates were made and what uses are appropriate given the data and methods that were used.

### Any additional comments on reporting **purpose**:


### Aim

Define the health indicator that is described and the populations for which estimates are made.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Should this item be in the reporting checklist?**

Include, as relevant, diagnostic criteria, target demographic group, and geographic areas for which estimates were made.

### Any additional comments on reporting the **aim** of a study:


Developing guidelines for reporting global health estimates: the GATHER project

Data inputs

Should each reporting item be a required checklist item, recommended checklist item, or should the reporting item be excluded from the checklist?

Tip: to navigate to another section of this survey, use the Question Index on the right of the screen or the previous/next button at the bottom of the page (you may need to scroll down).

Data search

Describe the data search strategy. Specify all databases searched.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In the medical field, a systematic data search usually includes a review of the published scientific literature. However, health data are often found in the grey literature, including health surveys, government reports, and databases maintained by international organizations and academic groups. Researchers should specify which databases were searched. They should also mention whether other search strategies were used, such as contacting professional networks for unpublished data. This allows readers to assess the comprehensiveness and completeness of the search. This reporting item refers to measures of the health indicator being estimated.

Any additional comments on reporting the data search strategy:

Inclusion/exclusion criteria

Explicitly define inclusion/exclusion criteria.
Consistent application of defined inclusion/exclusion criteria is a key characteristic of a systematic review. Information on these criteria will allow the reader to understand the quality of data used, as well as potentially informative data that were excluded. The authors should specify publication/data years included, languages, study designs, and demographic groups included, as relevant. This reporting item refers to measures of the health indicator being estimated.

Any additional comments on reporting the **inclusion and exclusion criteria**:

---

**Ad-hoc exclusions**

State whether any data sources which met inclusion criteria were eventually excluded. If so, identify the data sources and the reason(s) for exclusion.

Data sources which report implausible values are sometimes excluded on an ad-hoc basis. Ad-hoc exclusions should be explicitly identified so that the study can be replicated. This reporting item refers to measures of the health indicator being estimated.

Any additional comments on reporting **ad-hoc exclusions**:
**Data extraction (1)**

List and define all variables for which data were sought.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Should this item be in the reporting checklist?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adapted from PRISMA reporting guidelines for meta-analyses. This reporting item refers to characteristics of each data source, such as representativeness, data collection method, or demographic characteristics of the study population.

**Data extraction (2)**

Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate).

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Should this item be in the reporting checklist?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adapted from PRISMA reporting guidelines for meta-analyses. This reporting item refers to measures of the health indicator being estimated.

**Data extraction (3)**

Specify whether authors were contacted.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Should this item be in the reporting checklist?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Adapted from PRISMA reporting guidelines for meta-analyses. This reporting item refers to measures of the health indicator being estimated.
Any additional comments on reporting **data extraction**:

<table>
<thead>
<tr>
<th>Data description (1)</th>
<th>Identify and describe any types of input data that provide a potentially biased estimate of the quantity of interest.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Should this item be in the reporting checklist?</strong></td>
<td>require</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Input data may provide a biased estimate of the health indicator of interest. For example, case definitions or diagnostic criteria may vary between data sources. Readers may want to know whether estimates are informed by potentially biased input data. This reporting item refers to measures of the health indicator being estimated.</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Data description (2)</th>
<th>Provide a complete list of data sources used and their key characteristics. Should the reporting checklist require the following for every data source, if relevant:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>require</strong></td>
<td>recommend</td>
</tr>
<tr>
<td><strong>reference information</strong></td>
<td></td>
</tr>
<tr>
<td><strong>sampling strategy (e.g., simple random sample)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>population represented (e.g., subnational, occupation-based)</strong></td>
<td></td>
</tr>
</tbody>
</table>
This is vital information for understanding the basis of the estimates. For example, an estimated trend may be produced on the basis of a few studies of variable quality in a particular country. By defining the data used and their key characteristics, the user can understand the basis of the estimate for each population. This reporting item refers to measures of the health indicator being estimated.

Is there anything else that should be reported for every input data source?

Data description (3)

Provide a graphical representation of data availability in space and/or time.

In some cases, the list of data sources is lengthy. A graphical representation of data availability can illuminate whether there are geographical areas or time periods where data are particularly sparse.

Any other comments on reporting data description:
Covariates and other data inputs (1)

Describe and give the source of any other data inputs, such as covariates, used to generate health estimates.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?  

Often data that are not measures of the indicator being estimated are used to make estimates, such as covariates or model parameters in a natural history model. The sources for these data should be fully described.

Covariates and other data inputs (2)

For each data input, identify whether it met the relevant reporting guidelines (e.g., PRISMA guidelines for meta-analyses, these guidelines for health estimates, CONSORT guidelines for trial data).

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

Covariates or model parameters used in these studies may not adhere to the relevant reporting guidelines. These parameters may have an important influence on the health estimates.

Any additional comments on reporting covariates and other data inputs:

Replicability

Make all data inputs available. If confidential data were used, all remaining data should be made available and results of analyses using...
the subset of reported data should be given.

Should the reporting guidelines require:

<table>
<thead>
<tr>
<th></th>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>public access to all input data</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(excluding confidential data)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>public access to all covariate data?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>public access to all other model</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>parameters?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>provision of study results using</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>only public-access data inputs?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Researchers need access to data sets to promote replication and improvement in the methods. When some data are restricted access, the comparison of the full and available data sets let other analysts have a better understanding of the importance of the inclusion of the restricted data sets.

Any additional comments on replicability:

---
Developing guidelines for reporting global health estimates: the GATHER project

Analytic methods

Should each reporting item be a required checklist item, recommended checklist item, or should the reporting item be excluded from the checklist?

Tip: to navigate to another section of this survey, use the Question Index on the right of the screen or the previous/next button at the bottom of the page (you may need to scroll down).

Analytic overview (1)
Provide a lay description of the analytic methods, including what assumptions are implied by the methods.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

All users (technical and non-technical) should be able to understand the analysis. In particular, there may be assumptions embedded in models about how estimates are made, i.e., linear time trends or time trends based on covariates only. Some assumptions may particularly affect out-of-sample estimates. These should be clear to the user.

Analytic overview (2)
If the methods involves multiple analytic steps, provide a diagram showing model structure and/or analytic steps.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?
Studies often involve multiple analytic steps. These steps may be difficult to describe in text, while a flow diagram would be more understandable.

Any additional comments on reporting the analytic overview:

Data preparation
Describe any pre-processing of input data.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

Examples of data pre-processing include methods for calculation of population-level indicators from data on individuals, or pre-processing of covariate data. These steps should be described to allow for replication of the study.

Any additional comments on reporting the data preparation:

Model selection (1)
Describe the models that were considered during the model selection process and how the final model(s) were chosen. This includes providing a list of covariates evaluated, if covariates were used.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

This will give an indication of the quality of the model-building process.
Model selection (2)
Report how model performance was evaluated.

<table>
<thead>
<tr>
<th>Should this item be in the reporting checklist?</th>
</tr>
</thead>
<tbody>
<tr>
<td>require</td>
</tr>
</tbody>
</table>

Report whether and how model assumptions were checked. Report quantitative measures of goodness-of-fit and of cross-validation (if carried out). If k-fold cross validation is carried out, report the types of training and test sets used and the quantitative measures of goodness-of-fit used for evaluating model performance. Report if estimates were verified by comparisons between different models. These reports will give an indication of the thoroughness of the model-building process.

Model selection (3)
Provide a graphical representation of model fit, i.e., plots comparing model output to adjusted and/or unadjusted data inputs.

<table>
<thead>
<tr>
<th>Should this item be in the reporting checklist?</th>
</tr>
</thead>
<tbody>
<tr>
<td>require</td>
</tr>
</tbody>
</table>

Many users will want to understand how well the model fits the available data; a graphical representation is accessible to all users regardless of their statistical background. This can be included in an appendix.

Any additional comments on reporting model selection:

Methods description (1)
Provide a technical description of the model(s) finally used to calculate health estimates.
The technical description should be sufficient for replication. It should include mathematical formulae or statistical code for each analytical step and specification of all model parameters. The statistical package(s) used for the analysis should be named.

Methods description (2)
For each potentially biased input data type, report whether methods were developed to adjust the data or describe the reason for failing to adjust the data.

This is needed to understand whether/how use of potentially biased data could affect the final health estimates.

Any additional comments on reporting methods description:

Sensitivity analyses
Describe methods of sensitivity analyses (i.e. additional analyses) and findings, if done.

Adapted from PRISMA reporting guidelines for meta-analyses.
Any additional comments on reporting **sensitivity analyses**:

[Blank space]

**Uncertainty analyses**
Specify which sources of uncertainty are reflected in the uncertainty range.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>✔️</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

> Often uncertainty ranges reflect only a subset of sources of uncertainty. The sources of uncertainty which are reflected in the uncertainty range should be listed (e.g., sampling errors in input data, non-sampling errors in input data, uncertainty from multiple analytic steps, and model uncertainty may be reflected in final uncertainty ranges).

Any additional comments on reporting **uncertainty analyses**:

[Blank space]

**Replicability**
Make all statistical code available online.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td>❌</td>
<td>✗</td>
<td>✗</td>
<td>✗</td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

> Public access to code ensures that results are replicable. Reproducing an analysis from a description of methods in a paper is generally difficult and often impossible because there are too many details to be comprehensively described in the methods. For most researchers providing a general description of the approach with key assumptions, along with access to the code for more specific details would be easier to do.
Any additional comments on **replicability**: 

[Blank space for comments]
Results

Should each reporting item be a required checklist item, recommended checklist item, or should the reporting item be excluded from the checklist?

Tip: to navigate to another section of this survey, use the Question Index on the right of the screen or the previous/next button at the bottom of the page (you may need to scroll down).

Results
Provide health estimates in machine-readable format, disaggregated at the level used for reporting.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Should this item be in the reporting checklist?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This is important for straightforward use of estimates in other applications. This would not be necessary in a situation where only a few estimates are made, that is, for 5-10 world regions and one time period.

Any additional comments on reporting the analytic overview:

Uncertainty analysis
Report and interpret uncertainty ranges (e.g., 95% or 90% uncertainty intervals).

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Should this</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Uncertainty ranges should reflect the largest source(s) of uncertainty.

Any additional comments on reporting uncertainty:

---

### Evidence basis of results

**Categorize estimates for each population according to the amount and/or quality of input data from the same population that informed the estimates.**

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
</table>

**Should this item be in the reporting checklist?**

Most comprehensive global health estimates include estimates which are informed by variable amounts/quality of underlying input data. For example, there may be no data available for population A, one low-quality data source available for population B, and several high-quality data sources available for population C. Out-of-sample predictions (such as the one for population A) may be used to identify likely health priority areas in populations for which measured health data are not available, and can provide insight into probable trends in health indicators where trend data are not available or are inconsistent. However, it is important for users to recognize the limitations of these estimates. For example, out-of-sample health statistics are not a substitute for measurements when evaluating disease programs or health systems reforms. If possible, health estimates with missing or low-quality input data should be identified in graphical or tabular presentations of health estimates, e.g. by using a different color or font to identify these statistics.

Any additional comments on reporting evidence basis of results:

---
Time trends
If data support estimates of trends over time, provide estimates for multiple years (vs. a single point in time).

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Some groups make estimates for a specific year on a routine basis. Then others (including the press, general public and even scientists) have combined these multiple point estimates to produce trends which in fact were incorrect as both data and methods had changed between the different rounds of estimation. If all sets of estimates include trends over time, it would clarify to the user that different sets of estimates should not be compared.

Any additional comments on reporting trends over time:
Developing guidelines for reporting global health estimates: the GATHER project

Discussion

Should each reporting item be a required checklist item, recommended checklist item, or should the reporting item be excluded from the checklist?

Tip: to navigate to another section of this survey, use the Question Index on the right of the screen or the previous/next button at the bottom of the page (you may need to scroll down).

Comparison with prior estimates
If updating a previous set of estimates, the analyst should report the results of a decomposition analysis that explains changes.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Updated estimates may change due to updated datasets, methods changes, or both. The analyst should perform an analysis crossing datasets and methods to identify why estimates have changed. This will allow users to understand the cause of any differences between the two rounds of estimates.

Any additional comments on reporting **comparison with prior estimates**:

```

```

Limitations
Describe limitations in data inputs and analytic methods.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Important assumptions should be discussed, including the effects of analyst choices regarding data inclusion, selection of a model (or models) to generate health estimates, and methods for potentially biased data. Specifically identify major sources of uncertainty, including any that were not captured by the uncertainty range.

Any additional comments on reporting limitations:

Research needs
Describe studies that should be carried out to improve future estimates.

<table>
<thead>
<tr>
<th>require</th>
<th>recommend</th>
<th>exclude</th>
<th>no opinion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Should this item be in the reporting checklist?

Any additional comments on reporting research needs:
Developing guidelines for reporting global health estimates: the GATHER project

Suggestions

Can you think of anything else that authors should report when global health estimates are published?
Many thanks for your participation.

If you have any further comments, please send to Gretchen Stevens at stevensg@who.int

Print your answers.