GATHER:

Working group meeting to develop a standard reporting checklist for global health estimates

Lancet offices
London
18-19 February 2015

Contents

MEETING REPORT

Background

Global, regional, and country estimates of population health indicators are important for monitoring health and for guiding resource allocation. Because of the major data gaps and measurement challenges for health statistics, analytic methods are often needed to fill data gaps and maximize the usefulness of the available data. However, accurate interpretation and effective use of health estimates requires a detailed description of the input data on which estimates were based, including information on their quality, and a clear explanation of the methods used to derive estimates from the input data. To meet this challenge, a working group was convened, with the aim to define and promote good practice in reporting global health estimates. This aim will be met by developing Guidelines for Accurate and Transparent Health Estimates Reporting (GATHER), which will include a checklist of items to report whenever global health estimates are published. This report describes the proceedings of GATHER's face-to-face meeting in London 2015.

Prior to the meeting described in this report, the working group had developed draft documents outlining the scope of the guidelines and a comprehensive list of potential reporting items. Stakeholders, including analysts that generate global health estimates and their users, had been surveyed on the utility of the reporting items via an online survey. The objectives of the meeting were the following:

- Obtain advice and guidance from experts who had previously developed reporting guidelines in other areas
- Agree upon the scope of the reporting guidelines and a framework for selecting reporting items
- Agree on a list of items to be reported whenever health estimates are published
- Agree upon plans for publication and dissemination of the reporting guidelines.

In line with the meeting objectives, the meeting involved four main discussion areas about the reporting guidelines: (1) lessons from existing guidelines, (2) scope of the GHE reporting guidelines, (3) reporting items checklist, and (4) consultation and dissemination plans.

Lessons learned from existing reporting guidelines

The first session of the meeting involved a review of lessons learned from the development of other reporting guidelines. There were brief reviews provided by those who had played a central role in other guidelines:

Vivian Welch, PRISMA: An extension to the PRISMA guidelines was developed to address equity-focused analyses and was published in 2012. Key lessons included:

- They went through a similar process that GATHER is going through now. It is important to document why items are included, as there is not always empirical evidence as to why they are important, and journal editors will want to understand the reasoning behind what has been chosen.
- "Road testing" the guidelines before publication is crucial.
- Representation is important. In particular, the process should involve participants from low and middle income countries.
- Obtaining endorsements is critical, in particular from journals and funders. She suggested that we consider drafting text for journals to use in instructions for authors.
- Recommended that GATHER draft a longer Explanation and Elaboration document to accompany the guidelines.

Gary Collins, TRIPOD: Guidelines for the reporting of prediction models for prognosis and diagnosis were published in 2014. Key lessons included:

- The development of guidelines can be a long process. TRIPOD started in June 2011 and was only published in 2014. This process included four face-to-face meetings.
- Finding appropriate examples for the E&E document can be very time consuming, and those published in non-open access articles require extra work to obtain copyright permissions.
- For TRIPOD, they have conducted an assessment of baseline adherence to the guidelines, and will repeat it in 5 years to see if reporting has improved.

Ana Marusic, EQUATOR: The EQUATOR network is an umbrella organization that seeks to improve reporting and quality of health research, and houses a library of reporting guidelines.

- Noted that the difference between reporting and conduct is never clear and the goal of reporting guidelines should ultimately be to improve the quality of reporting and the quality of research. Reporting checklists can serve an educational role in terms of reviewing best practices.
- It is important to establish a plan for evaluating the use and effectiveness of guidelines once implemented.

Paul Simpson, PLOS Medicine:

- Urged the group to consider the importance of buy-in from journal editors.
- Noted that for many journals, the editors are volunteer medical editors, rather than full-time, paid professional editors. This limits the amount of time they have to spend reviewing and implementing new guidelines. Given this, the EQUATOR network can play a valuable role as it serves as a home for the guidelines.
- Highlighted the issue that many authors view checklists as a final hurdle to publication, rather than a helpful input to the writing and publication process.
- PLOS has recently implemented an open access data policy, which was challenging.
- Noted that good reviewers tend to use guidelines to support their review.

John Grove gave a brief update on the new Bill and Melinda Gates Foundation open access policy, which requires the immediate provision of open access and re-analysable data for projects they fund. There will be a two year ramp up of the policy, to ensure it is done properly, and the foundation will fund the costs of open access. It is expected to be fully implemented by 2017.

Scope of GATHER

In conference calls leading up to the meeting, the working group had suggested a set of health indicators that would fall within the scope of the reporting guidelines based on the following definition:

"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, including health behaviors, health exposures, and intervention coverage."

During the meeting, it was noted that this scope comprises three main categories: (1) measures of health outcomes, (2) measures of health risk factors, and (3) measures of intervention coverage. Several meeting participants expressed concern that indicators of intervention coverage may involve additional methodological considerations that would not necessarily be captured by a general set of

reporting guidelines, in particular with respect to the challenges of appropriately defining and measuring the quality of a health service in conjunction with its coverage level. While it was noted that lack of a particular health service can be considered a health risk factor, the group reached consensus that the scope of the guidelines should be more limited than originally proposed. Thus, indicators of health status, health behaviors and health exposure would be included, but indicators of intervention coverage would not. It was suggested that indicators of intervention coverage could be added in the future as an extension, if deemed necessary and if health services researchers are involved in the discussions. Additionally, the text of the statement document would note that these reporting guidelines could be of use to those publishing estimates of health services coverage.

Other important points for the discussion on the scope of the guidelines included:

- Future forecasts of health indicators are not included as they are "outside the data".
- A summary statistic from a country health survey would not fall under these reporting guidelines. However, this raised the question about direct reporting of a country's VR data. It was agreed that there may be a gray area which need not be delineated at this time. For now, the group agreed to focus on what is definitely excluded (single data source from a country) vs. what is clearly included (synthesis and adjustment of multiple data sources).
- There was also discussion about the extent to which the guidelines should aim to dictate the methods for global health estimates versus the reporting of global health estimates. There was consensus among the group that these guidelines should be for reporting, not conduct, following the approach taken by other guidelines like PRISMA.
- WHO staff noted that they hoped these guidelines would be adopted for WHO reports. They should therefore be broad enough to be applicable to reports and journal articles.

Reporting items checklist

The majority of the meeting involved discussing, point-by-point, the items considered for the reporting guidelines checklist. This draft checklist was previously compiled by the working group. The draft checklist was meant to be exhaustive, and no suggested items were excluded. In the meeting it was agreed that the group should aim to reduce the draft list to a more concise list of items considered essential. There was a brief discussion of a two-tiered checklist list, with some items listed as "required" and others listed as "recommended". The group decided that the list should only include "required" (labelled "essential"), with additional text in the statement document providing suggestions for less essential, but potentially useful, items, as well as examples and key issues pertaining to essential items included in the final checklist.

Prior to the meeting, an online survey was conducted to solicit feedback from GHE users and producers about each of the 44 potential checklist items and sub-items. Survey responses from 118 individuals were summarized for the meeting. For each item, bar charts were shown depicting the number of survey respondents who thought a reporting item should be required, recommended, neutral/no opinion, or excluded, and all submitted comments from the survey were provided to the working group. Responses to the survey were not meant to be definitive, but rather to serve as a guide to the working group about views on each reporting item in the wider GHE community. For all items, the majority of respondents felt an item should be either required or recommended.

The draft checklist grouped items into 5 domains: (i) context and design, (ii) data inputs, (iii) analytic methods, (iv) results and (v) discussion. Below is a summary of the key points discussed for each of these domains. The final proposed checklist is given at the end of this section:

(i) Context and design

- The group agreed that an abstract should not be required, as it is typically included in most publications but with formats that are so different that it would not be feasible to specify what the abstract should include with respect to reporting. Specifically, the group could not agree upon any single reporting item that was deemed essential to include in an abstract.
- There was consensus that the following information should be reported: indicator definition, population (age, sex, geographic region) and time period. This was labelled "aim" on the draft checklist, but should be renamed.

(ii) Data inputs

- The group discussed reporting data inputs in detail, in particular a tabular summary listing key variables for each data source, as well as providing non-confidential data open access along with the publication of estimates.
- Types of data inputs were also discussed, as not all may have the same reporting requirements. For example, if an analyst is estimating the prevalence of disease with a regression model, survey-based estimates of prevalence extracted from survey reports likely require more reporting than the covariates used in the regression equation. It was agreed that data inputs be placed in two categories: data inputs that are synthesized for the purposes of the study (e.g., survey estimates of diabetes prevalence for a study aiming to estimate national diabetes prevalence) and data taken from existing sources (e.g., GDP series from the World Bank in a study using GDP as a covariate in a regression).
- For data inputs that are synthesized for the purposes of the study, there should be a reporting item that lists each data source and provides the key characteristics for each source. This table would include, for each data source, its: sampling strategy, population, year of data collection, sex and age range, diagnostic criteria or measurement method, and sample size. It would also include column(s) for other major, know important sources of bias or incomparability across sources (in most cases, these are expected to already be covered by age, sex, population, diagnostic criteria or measurement method). It was noted that not all columns would be needed if they form part of inclusion criteria, e.g., if all data sources use the same measurement method it would not be useful to specify it for every data source.
- For the above table, it was noted that in some rare cases, sources could not be named due to genuine concern about serious risk of harm to those who collected or maintain the data.
- For data inputs that are not synthesized for purposes of the study, these inputs should be described and their sources provided. Examples of these inputs may include: population sizes, covariates and parameter values for natural history models.
- The group agreed that it was not essential for authors to indicate whether all input data met relevant reporting guidelines.
- The group agreed that a reporting item called "Data transparency" should be included, under which authors make all data inputs available, with the exception of confidential data. Contact names or institutions should be given for confidential data.
- The group further agreed that the checklist should not include an item requiring that authors conduct a sensitivity analysis in which they re-estimate all outcomes of interest using only the publically available subset of data, in cases where confidential data are included in the data set.

(iii) Analytic methods

- There was extensive discussion about how structured the reporting items should be for reporting on analytic methods. The consensus of the group was that the checklist should not be too prescriptive, as each study has different characteristics. It was finally agreed that much of the description of analytic methods be condensed in one reporting item that would provide suggested components, namely: data cleaning, data pre-processing, data adjustments and weighting of data sources, and technical description of the model(s). There would also be a suggestion that the authors avoid jargon and provide explanations for non-specialists.

- After debating several options, the group agreed on one reporting item for "Model building", which would summarize how models were developed and selected and a second item for "Model evaluation", which would summarize the methods used to evaluate model performance and, if conducted, sensitivity analyses.
- There was extensive discussion about a reporting item for uncertainty. The concern was raised that in some cases, meaningful uncertainty intervals cannot be calculated as the non-sampling error is large and unknown. The final consensus was that a reporting item for uncertainty should be worded as: "what sources of uncertainty are, and are not, reflected in the uncertainty intervals". Furthermore, the statement document should provide an overview of major sources of uncertainty analysts should consider, and also suggest that authors be explicit about how uncertainty is propagated through the estimation process.
- There was extensive discussion about a reporting item that would require that all computer code used in the estimates be provided along with publication. Some key points raised included:
    - The time required to prepare open access statistical code is a burden on research groups, and a practical barrier to meeting this requirement.
    - The IHME's GBD estimation process is too complicated to fully meet this requirement.
    - To find middle ground, it was proposed that authors need only provide the component of their code at the core of the estimation process, e.g., R code for a statistical model, as opposed to providing a fully functional piece of "software" that reproduces the entire data cleaning, pre-processing, adjustment and analysis steps with the push of a button.
    - A compromise was proposed that authors state where the code can be obtained, and then if they chose to not make it available, they must state that explicitly in the publication. This would follow the approach used by *Nature*, which requires access to all code and sources for editors, reviewers and readers, with any restrictions stated explicitly in the submitted manuscript.
    - The final agreement of the group was that there be a reporting item that states where the code can be obtained, with a recommendation that the code be posted as open access. However, IHME expressed reservations about this reporting item due to the extensive number of steps involved in the GBD estimation process.

(iv) Results

- It was agreed that all results reported in a publication should also be made available in an accessible data format, e.g., a CSV or excel file with country-years of obesity prevalence should be made available if the paper presents trends in national obesity.
- It was agreed that authors should address uncertainty in the results, with a suggestion of presenting confidence intervals for estimated quantities. If the authors cannot compute meaningful confidence intervals, they should discuss the reason for this.
- There was general agreement on the need for several specific aspects of communicating the results of global health estimates, however none of these related to issues that were universal across all types of estimates and were therefore not deemed essential. The group suggested that some of these topics be discussed in the statement document. These topics included: graphical representation of final results along with (un)adjusted data inputs, estimated regression coefficients, and flagging estimates based on underlying data availability. It was suggested that the group consider creating a generic results reporting item like "Clearly report results" which would provide space for listing some of these points.

(v) Discussion

- It was agreed that there should be a reporting item for limitations of the analysis, and that the text of the statement document should summarize 5-6 categories of limitations that authors should consider discussing.

The group also discussed three other items that sometimes appear on reporting checklists: conflict of interest, contact information and funding source. It was noted that “conflict of interest” has negative connotation and that “Declaration of interests” would be more accurate. It was unclear how a declaration of interest, or contact information, would apply for WHO reports, as sometimes they have no explicit authors. The group decided that reporting the funding source was applicable in nearly all settings and should be a reporting item, with contact information and reference information as suggested items.

Table 1 contains the components of the checklist that were agreed upon by the working group at the end of the meeting, although the wording of the items must still be finalized.

Table 1.List of the GATHER reporting items. Wording of the items remains to be finalized.

Context and Design
1. Definition of the indicator(s) estimated, populations (including age, sex, geographic region) for which estimates are made, and time period(s) for which estimates were made.
Data Inputs
For data that were synthesized in this study:
2. Describe the data identification and access strategy
3. Define the inclusion/exclusion criteria, including ad-hoc exclusions (e.g., outliers)
4. Identify and describe types of input data that have potentially important biases
5. Provide key characteristics of input data sources - Table with information for each study, including: reference information or contact name/institution, sampling strategy, population represented, year of data collection, sex and age range, diagnostic criteria or measurement method, sample size - Include column(s) for other major, known important sources of bias for which analytic methods were developed (most of these are already covered by age, sex, population, diagnostic criteria or measurement method)
For data inputs that were used without modification:
6. Describe and give sources for other inputs (for data not synthesized for study) - examples: population sizes, covariates, parameter values in natural history models
For all data inputs:
7. Make all data inputs available, with the exception for confidential data, for which contact names/institutions are made available
Analytic Methods
8. Provide a [conceptual] overview of analytic steps (a diagram may be useful)
9. Detailed description of analytic steps, including mathematical formulas (a) data cleaning, if applicable (b) data pre-processing, if applicable (c) adjustments + weighting of data sources, if applicable (d) technical description of model(s), including every analytic step *suggestion to avoid jargon and provide explanation for non-specialists
10. Model building (model development process and model selection)
11. Model performance (including sensitivity analysis, if relevant)
12. Uncertainty - What sources of uncertainty are, and are not, reflected in the uncertainty intervals? Describe methods for calculating uncertainty in terms that can be understood by non-specialist.
13. State where computer/statistical code can be accessed - The 'gold standard' – a link to code which may be used off-the-shelf together with input data – will be described, and then alternative ways to fulfil this reporting guideline, allowing for different resourcing levels and complexity of code, will be given
Results
14. Clearly describe results. This item includes a series of suggested 'best practices' for describing results in a way that is accessible to all users.
15. Provide published estimates in electronic data format (e.g., CSV file)
16. Address uncertainty, with recommendation confidence intervals be given. Discuss why uncertainty was not calculated, if not calculated.
Discussion
17. Discuss estimates in view of existing evidence (e.g., other estimates), and if updating previous set of estimates, discuss reasons for changes in estimates
18. Discuss limitations of the estimates
19. List the funding sources for the work

Consultation and dissemination

The final session of the meeting focused on next steps, in particular (i) plans for drafting documents, (ii) inviting wider feedback, and (iii) dissemination.

(i) Three primary outputs are anticipated: first, a checklist of reporting items (maximum 2 pages); second, a statement, including the checklist, for publication; and third, an explanatory document giving evidence and justification for each item on the checklist. The drafting plan for the checklist and statement is as follows:

1. WHO staff led by Gretchen Stevens will write a draft of the statement document and the reporting items checklist.
2. This draft will be circulated to members of the working group for comment (anticipated timeline: March). During this step, the working group members will suggest examples of good practice for each reporting item.
3. After WHO staff incorporate comments from the working group, the final draft will be circulated for larger consultation (anticipated timeline: April/May).

It was also agreed that an Explanation and Elaboration (E&E) should be developed simultaneously, which would include examples of best practice for each of the reporting guidelines. It was noted that E&E documents that contain open access tables and figures are easier to publish as there is no need to obtain copyright permission. However, the group recommended that the E&E not be restricted to open access to ensure a wide sampling of GHE publications in the examples.

Other outputs, including a peer review checklist and systematic review of current reporting, were discussed but not prioritized at this time.

(ii) It was noted that much wider consultation is needed for the proposed reporting guidelines to ensure buy-in from the global health community. In particular, it was suggested that more representation from low-and-middle income countries was needed, as was comment from other UN agencies that produce global health estimates. The following organizations were listed: UNICEF, UN population division, and other UN agencies involved in global health estimates, INDEPTH network, International Epidemiological Association, Public Health Foundation of India, WHO focal points for estimates, GBD expert community and Independent Advisory committee, national and European CDCs, Ideas country offices (via Gates), DFiD evidence for action group's country offices; several individuals were mentioned, including David Evans, Dean Jamison and Peter Smith. Working group members will collaborate on finalizing the list of partners to consult and reaching out to these partners. These groups will be asked to comment simply on whether they find the guidelines useful for addressing mistrust of estimates and whether there is anything critical to add. The group also discussed circulating the draft for initial comment from journals that are heavily involved in publishing global health estimates. Richard Horton agreed to select and contact 5-10 key journals on behalf of the group.

(iii) There was tentative agreement for *The Lancet* and *PLOS Medicine* to jointly publish the reporting guidelines, ideally in June to link to the Measurement Summit in Washington. Publishing in a wider set of journals was discussed, but it was noted that this creates copyright problems and is, in general, not a good use of time. If the guidelines are published open access, other journals can publish editorials endorsing the guidelines. There was agreement we needed to obtain a Creative Commons license for the guidelines so they would be fully open access and reusable. It was agreed that a living E&E document would be available online upon publication of the guidelines for reference. It was noted that a word version of the checklist should be made available at the same time.

It was agreed that once the guidelines are close to being finalized, they should also be circulated more widely to speciality journals that publish infrequently publish global health estimates and may view these guidelines as useful for structuring review processes. For example, the guidelines can be shared via the ICMJE and WAME listservs.

Towards the close of the meeting, the group discussed “test driving” the guidelines. Several members of the group pledged to apply the guidelines to one paper they are currently working on, as a test of their feasibility: Leontine Alkema, Theo Vos, Igor Rudan, and Joy Lawn. Joy Lawn and Ana Marusic agreed to head up dissemination of the guidelines.

Annex A. Agenda

February 18, 2015

		Presenters/Chairs
9:00 – 12:00	**Plan of work**	
9:00	Welcome and introductions	Richard Horton / Colin Mathers
9:15	Background, method of work, and results of online survey	Gretchen Stevens
9:40	Key lessons and experiences: PRISMA-E guidelines	Vivian Welch
	Key lessons and experiences: TRIPOD guidelines	Gary Collins
	Reporting guidelines: EQUATOR perspective	Ana Marusic
	Reporting guidelines: journal editors' perspectives	Paul Simpson
	BMGF open-access policy	John Grove
10:20	Discussion / agreement on method of work	All
10:30	Coffee break	
11:00 -12:00		
11:00	Scope and technical terms: discussion and adoption by the group	All
12:00	Lunch break	
13:00 -- 15:30	**Context and design / Data inputs**	
13:30	Review of reporting items	Chair: Bob Black
15:30	Coffee break	
16:00 – 18:00	**Data inputs**	
16:00	Review of reporting items	Chair: John Grove
19:00	Group Dinner: Dishoom Shoreditch 7 Boundary Street, London E2 7JE	

February 19, 2015

9:00 – 10:30	**Analytic methods**	
9:00	Review of reporting items	Chair: Gary Collins
10:30	Coffee break	
11:00-12:30	**Results**	
11:00	Review of reporting items	Chair: Paul Simpson
12:30	Lunch break	
14:00-16:00	**Discussion and next steps**	
14:00	Review of reporting items	Chair: Richard Horton
15:30	Plans for writing and publication	All
16:00	Meeting end	

Annex B. List of Participants

Leontine ALKEMA National University of Singapore Faculty of Science	alkema@nus.edu.sg
Robert BLACK Institute of International Programs at Johns Hopkins Bloomberg School of Public Health	rblack@jhsph.edu
Gary COLLINS University of Oxford, Centre for Statistics in Medicine (CSM)	gary.collins@csm.ox.ac.uk
Majid EZZATI Imperial College London	majid.ezzati@imperial.ac.uk
John GROVE GATES Foundation	John.Grove@gatesfoundation.org
Dan HOGAN World Health Organization	HoganD@who.int
Richard HORTON The Lancet	Richard.Horton@lancet.com
Joy LAWN London School of Hygiene & Tropical Medicine	joy.lawn@lshtm.ac.uk
Ana MARUSIC University of Split, School of Medicine	ana.marusic@gmail.com
Colin MATHERS World Health Organization	MathersC@who.int
Igor RUDAN University of Edinburgh, Centre for Population Health Sciences	irudan@hotmail.com
Paul SIMPSON PLOS Medicine	psimpson@plos.org
Gretchen STEVENS World Health Organization	StevensG@who.int
Theo VOS University of Washington, Institute for Health Metrics and Evaluation	tvos@uw.edu
Vivian WELCH University of Ottawa Faculty of Medicine, Centre for Global Health	Vivian.Welch@uottawa.ca