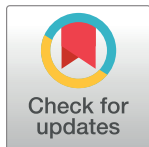


GUIDELINES AND GUIDANCE

Recommended reporting items for epidemic forecasting and prediction research: The EPIFORGE 2020 guidelines

Simon Pollett¹, Michael A. Johansson², Nicholas G. Reich³, David Brett-Major⁴, Sara Y. Del Valle⁵, Srinivasan Venkatramanan⁶, Rachel Lowe^{7,8}, Travis Porco⁹, Irina Maljkovic Berry¹, Alina Deshpande⁵, Moritz U. G. Kraemer¹⁰, David L. Blazes¹¹, Wirichada Pan-ngum¹², Alessandro Vespigiani¹³, Suzanne E. Mate¹, Sheetal P. Silal^{14,15}, Sasikiran Kandula¹⁶, Rachel Sippy¹⁷, Talia M. Quandelacy², Jeffrey J. Morgan¹⁸, Jacob Ball¹⁹, Lindsay C. Morton^{20,21}, Benjamin M. Althouse^{22,23,24}, Julie Pavlin²⁵, Wilbert van Panhuis²⁶, Steven Riley²⁷, Matthew Biggerstaff²⁸, Cecile Viboud²⁹, Oliver Brady³⁰, Caitlin Rivers³¹*



OPEN ACCESS

Citation: Pollett S, Johansson MA, Reich NG, Brett-Major D, Del Valle SY, Venkatramanan S, et al. (2021) Recommended reporting items for epidemic forecasting and prediction research: The EPIFORGE 2020 guidelines. *PLoS Med* 18(10): e1003793. <https://doi.org/10.1371/journal.pmed.1003793>

Published: October 19, 2021

Copyright: This is an open access article, free of all copyright, and may be freely reproduced, distributed, transmitted, modified, built upon, or otherwise used by anyone for any lawful purpose. The work is made available under the [Creative Commons CC0](https://creativecommons.org/licenses/by/4.0/) public domain dedication.

Funding: We appreciate the role of the Outbreak Science and Model Implementation Working Group in developing this initiative, and the Johns Hopkins Center for Health Security for hosting the face-to-face consensus meeting and conducting the electronic Delphi process. We would also like to thank the MIDAS Coordination Center and the National Institutes of General Medical Sciences (NIGMS 1U24GM132013) for supporting travel to the face-to-face consensus meeting by members of the Working Group. NGR was supported by the National Institutes of General Medical Sciences (R35GM119582). Travel for SV was supported by the National Institutes of General Medical Sciences (1U24GM132013-01). BMA was supported by Bill & Melinda Gates through the Global Good Fund. RL was funded by a Royal Society Dorothy

1 Walter Reed Army Institute of Research, Silver Spring, Maryland, United States of America, **2** Division of Vector-Borne Diseases, Centers for Disease Control & Prevention, San Juan, Puerto Rico, United States of America, **3** University of Massachusetts—Amherst, School of Public Health and Health Sciences, Amherst, Massachusetts, United States of America, **4** University of Nebraska Medical Center, Omaha, Nebraska, United States of America, **5** Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America, **6** Biocomplexity Institute and Initiative, University of Virginia, Charlottesville, Virginia, United States of America, **7** Centre for Mathematical Modelling of Infectious Diseases and Centre on Climate Change and Planetary Health, London School of Hygiene & Tropical Medicine, London, United Kingdom, **8** Barcelona Institute for Global Health, Barcelona, Spain, **9** University of California at San Francisco, San Francisco, California, United States of America, **10** Department of Zoology, University of Oxford, Oxford, United Kingdom, **11** Bill and Melinda Gates Foundation, Seattle, Washington, United States of America, **12** Mahidol-Oxford Tropical Medicine Research Unit and Department of Tropical Hygiene, Mahidol University, Thailand, **13** Network Science Institute, Northeastern University, Boston, Massachusetts, United States of America, **14** Modelling and Simulation Hub, Africa, Department of Statistical Sciences, University of Cape Town, Cape Town, South Africa, **15** Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom, **16** Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York City, New York, United States of America, **17** Institute for Global Health and Translational Science, State University of New York Upstate Medical University, Syracuse, New York, United States of America, **18** Catholic University of America, Washington, DC, United States of America, **19** U.S. Army Public Health Center, Edgewood, Maryland, United States of America, **20** Armed Forces Health Surveillance Division, Global Emerging Infections Surveillance, Silver Spring, Maryland, United States of America, **21** George Washington University, Milken Institute School of Public Health, Washington, DC, United States of America, **22** University of Washington, Seattle, Washington, United States of America, **23** Institute for Disease Modeling, Bellevue, Washington, United States of America, **24** New Mexico State University, Las Cruces, New Mexico, United States of America, **25** National Academies of Sciences, Engineering, and Medicine, Washington, DC, United States of America, **26** University of Pittsburgh Graduate School of Public Health, Pittsburgh, Pennsylvania, United States of America, **27** MRC Centre for Global Infectious Disease Analysis, School of Public Health, Imperial College, London, United Kingdom, **28** Influenza Division, Centers for Disease Control & Prevention, Atlanta, Georgia, United States of America, **29** Fogarty International Center, National Institutes for Health, Bethesda, Maryland, United States of America, **30** Centre for the Mathematical Modelling of Infectious Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom, **31** Johns Hopkins Center for Health Security, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland, United States of America

* crivers6@jhu.edu

Hodgkin Fellowship. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Abstract

Background

The importance of infectious disease epidemic forecasting and prediction research is underscored by decades of communicable disease outbreaks, including COVID-19. Unlike other fields of medical research, such as clinical trials and systematic reviews, no reporting guidelines exist for reporting epidemic forecasting and prediction research despite their utility. We therefore developed the EPIFORGE checklist, a guideline for standardized reporting of epidemic forecasting research.

Methods and findings

We developed this checklist using a best-practice process for development of reporting guidelines, involving a Delphi process and broad consultation with an international panel of infectious disease modelers and model end users. The objectives of these guidelines are to improve the consistency, reproducibility, comparability, and quality of epidemic forecasting reporting. The guidelines are not designed to advise scientists on how to perform epidemic forecasting and prediction research, but rather to serve as a standard for reporting critical methodological details of such studies.

Conclusions

These guidelines have been submitted to the EQUATOR network, in addition to hosting by other dedicated webpages to facilitate feedback and journal endorsement.

Introduction

The importance of infectious disease epidemic forecasting and prediction research is underscored across decades of communicable disease outbreaks. Epidemic forecasts are valuable for seasonal pathogens, for example, influenza and dengue [1–3], in addition to international health public emergencies and other epidemics such as the Zika, chikungunya, and Ebola virus epidemics [4–9]. Most recently, the Coronavirus Disease 2019 (COVID-19) pandemic has illustrated the importance of robust, transparent epidemic forecasting and prediction research for risk communication, decision-making, preparedness, and response [10,11]. Arguably, predictions form an essential part of the scientific method itself [12].

Other fields of medical research, such as clinical trials and systematic reviews, have widely used study reporting checklists, for example, the CONSORT and PRISMA guidelines [13]. Such checklists improve the interpretation, evaluation, and reproduction by other scientists and stakeholders, including public health decision-makers, journal editors, and journal reviewers. Indeed, many journals mandate that reporting checklists are completed prior to manuscript submission and publication, which has led to demonstrable improvements in study reporting [14,15]. Although principles for policy-driven communication of models for neglected tropical disease programs have been discussed [16], a recent systematic review noted no reporting guidelines exist specifically for epidemic forecasting and prediction research [17]. The need for epidemic forecasting reporting guidelines is underscored by a review of Zika forecasting and prediction research, which noted methodological reproducibility, accessibility, and incorporation of uncertainty in these published predictions varied [8].

To address this gap, we developed the EPIFORGE checklist, the first known set of epidemic forecasting reporting guidelines. This checklist was developed through a well-established process for developing guidelines for research reporting, involving a Delphi process and broad consultation with an international panel of infectious disease modelers and model end users [18,19]. The objectives of these guidelines are to improve the consistency, reproducibility, comparability, and quality of epidemic forecasting reporting. Here, we describe our guidelines development process and the resulting checklist. The EPIFORGE checklist is not designed to advise scientists how to perform epidemic forecasting and prediction research, but rather serve as a set of standards to ensure critical aspects of these studies are reported in a standardized way.

Methods

We followed health research reporting guideline development best practice as outlined in the EQUATOR toolkit and by Moher and colleagues [18,19] and summarized in full in the Supporting information (S1 Text). Briefly, The EPIFORGE guideline concept was registered at the EQUATOR network, and a steering committee ($n = 6$) formed to develop a guideline development protocol. Members from this steering committee had already identified a case study that prompted the need for EPIFORGE and conducted a systematic review to ensure no epidemic forecasting reporting guideline existed [17]. The EPIFORGE steering committee formulated an initial draft checklist of 20 reporting items during 2 teleconferences. This draft checklist was the input for an iterative Delphi consensus process. A total of 69 Delphi panelists were invited, and 46 participated in this process (S1 Table). During 3 initial rounds of Delphi consultations via email (September, October, and December 2019), panelists graded each checklist item on a scale of 1 through 10 (a score of 1 was defined as “not important,” and a score of 10 was defined as “very important”), with an emphasis on voting based on the concept of the item (rather than the wording). Checklist items with a mean score ≥ 8 were retained for the final reporting checklist, items with a mean score < 5 were dropped, and items with a mean score 5 to 7 were kept for further discussion at a final face-to-face consensus meeting (January 2020). Additional items were added by Delphi participants during the first 2 email Delphi rounds.

Results

Table 1 presents the final consensus checklist items, including reporting elements on study goals, data sources, model characteristics and assumptions, model evaluation, and study generalizability. Below, we elaborate and explain each item:

A. Overall study description and goals

Item 1: Describe the study as forecast or prediction research in at least the title or abstract. These guidelines primarily refer to forecasting research, but the principles are applicable to prediction research more broadly and can be used for these other study types. Forecasting research has been defined as research that “typically offers quantitative statements about an event, outcome, or trend that has not yet been observed, conditional on data that has been observed,” whereas prediction research is broader and has been defined as a field, which “may refer to models that examine the mechanistic drivers of epidemiological characteristics . . . as well as studies that estimate epidemiological characteristics with inherent forecasting value, such as R_0 ” [17]. As the 2 terms may be conflated by studies, we recommend that the study is described as a forecast or prediction research in at least the title or abstract (for example, [20–22]). While limiting to the terms “forecast,” “forecasting,” or “prediction” may be too restrictive, we believe that limiting the number of terms is important to enable findability (accurate

Table 1. EPIFORGE 2020 checklist.

Section of manuscript	#	Checklist item	Reported on page ^a
Title/Abstract	1	Describe the study as forecast or prediction research in at least the title or abstract	
Introduction	2	Define the purpose of study and forecasting targets	
Methods	3	Fully document the methods	
Methods	4	Identify whether the forecast was performed prospectively, in real time, and/or retrospectively	
Methods	5	Explicitly describe the origin of input source data, with references	
Methods	6	Provide source data with publication, or document reasons as to why this was not possible	
Methods	7	Describe input data processing procedures in detail	
Methods	8	State and describe the model type, and document model assumptions, including references	
Methods	9	Make the model code available, or document the reasons why this was not possible	
Methods	10	Describe the model validation, and justify the approach	
Methods	11	Describe the forecast accuracy evaluation method used, with justification	
Methods	12	Where possible, compare model results to a benchmark or other comparator model, with justification of comparator choice	
Methods	13	Describe the forecast horizon, with justification of its length	
Results	14	Present and explain uncertainty of forecasting results	
Results ^b	15	Briefly summarize the results in nontechnical terms, including a nontechnical interpretation of forecast uncertainty	
Results	16	If results are published as a data object, encourage a time-stamped version number	
Discussion	17	Describe the weaknesses of the forecast, including weaknesses specific to data quality and methods	
Discussion	18	If the research is applicable to a specific epidemic, comment on its potential implications and impact for public health action and decision-making	
Discussion	19	If the research is applicable to a specific epidemic, comment on how generalizable it may be across populations	

^aThis column refers to where key reporting considerations are included in a manuscript.

^bA break-out box may be a preferred location.

<https://doi.org/10.1371/journal.pmed.1003793.t001>

returns on searches) in the literature and may assist in standardizing nomenclature across the field.

Epidemic forecasts may be conflated with projections, simulations, or scenario analyses [4,10,20], which may or may not be fundamentally different in nature. Here, we refer to forecasts and predictions as predictions of what *will* happen. Other valuable research focuses on projections, simulations, or scenario analyses, which can be framed as “what if” scenarios, i.e., what would happen *conditionally* under certain conditions or assumptions (including, for example, assumptions of no interventions or no control of an epidemic, or assumptions of long-term seasonality and varied population immunity) [23,24]. Forecasts often refer to more shorter-term predictions [23]. Many of the reporting principles in this guideline may also be useful for projections, simulations, or scenario analyses.

Item 2: Define the purpose of study and forecasting targets. Clearly identifying the research objectives is a fundamental element of any scientific study and is a feature of many other research reporting guidelines [19]. We recommend that forecasting targets (i.e., each specific observable outcome being forecasted such as a 2-week-ahead incidence, peak week,

observation of at least 1 case) should be defined in the introduction section, and, ideally, also in the abstract (for example, [21,22,25]).

Item 3: Fully document the methods. Methods documentation is essential to any scientific study and follows general best practice for the reporting of other research study types [19]. We recommend that forecasting methods should include a full description of the model that enables reproducibility, the method of fitting parameters to data (for example, maximum likelihood with function if nonstandard, Bayesian methods), and—where relevant—underlying epidemic model assumptions (see also **Item 8**). For example, [20].

Item 4: Identify whether the forecast was performed prospectively, in real time, and/or retrospectively. We recommend that it is identified whether the forecast was performed prospectively, in real time, and/or retrospectively (for example, [22]). This item is necessary for interpreting results of forecasting accuracy and may aid in determining whether authors were blinded to a hold-out set (out-of-sample set) of data used for any model validations. See also **Item 16** for recommendations on time-stamping the results of forecasts.

B. Data description

Item 5: Explicitly describe the origin of input source data, with references. We recommend that the origin of the input source data is provided (for example, [21,26]). This item is essential for study reproducibility and is a minimum requirement for any manuscript, even if full study data cannot be publicly shared (see **Item 6**). For all data types—including laboratory assay, case counts, demographic data, and nontraditional data streams (for example, internet event-based data signals)—the authors should include sufficient references to be able to identify the input data [27] and ideally a persistent and unique identifier that resolves to the (meta) data (for example, [6]).

Item 6: Provide source data with publication, or document reasons as to why this was not possible. We recommend that source data are made available. Provision of source data improves forecast reproducibility. Sharing of source data used in forecasts (for example, [1]) facilitates other complementary studies, including those which may independently validate forecasts and methods. Limitations on data sharing during epidemics is a known challenge [28]. We are aware of efforts to establish codes of conduct for data sharing during public health emergencies [29] but recognize the wide range of logistical and other barriers to data sharing during outbreaks [28,30]. Therefore, we suggest at a minimum reporting of the reasons for not providing source data with forecast publication. Several major biomedical journals now routinely require authors to provide deidentified data [31]. When data are provided, we recommend inclusion of a data dictionary and/or structured metadata in a standardized format.

Item 7: Describe input data processing procedures in detail. We recommend that the input data processing procedures are described in detail (for example, [21,22]). This is an important feature for study reproducibility. Preprocessing procedures may include recoding and imputation of missing observations, identification and management of extreme outliers and influential data points, and functional transformations such as data normalization. Provision of data preprocessing code may also be useful.

C. Model characteristics

Item 8: State and describe the model type, and document model assumptions, including references. We recommend that the model type is stated and described, and the model assumptions documented (for example, [21,22]). This is critical for study reproducibility, and it allows interpretation of model output in the context of any assumptions presented. Describing model parameter values and assumptions, with references, further allows other researchers

to use cited parameter values in their own work (after careful consideration), and this may expedite forecasting efforts in a public health emergency. For an ongoing epidemic, if the model makes specific assumptions about current and future interventions and their impact, they need to be stated with appropriate justification. Model types may include mechanistic or statistical representation of disease transmission; models may also be classified as stochastic or deterministic models. These are just some of several classifications of model types [32]. We do not propose a categorization scheme for model types in these guidelines due to the wide range of model type nomenclature that is often heterogeneously used by modelers. Developing such a schema could be the subject of future research.

Item 9: Make the model code available, or document the reasons why this is not possible. We recommend that model code is made available. Providing model code improves research reproducibility, especially if accompanied by documentation, and may facilitate the rapid conduct of other studies addressing the same or similar study question(s), especially during a public health emergency. Some forecasting studies already have provided model code during public health emergencies of international concern [6,33]. Infectious disease modelers have also made the point that publication of model code may permit direct comparisons of model performance in real time by external groups [7]. We emphasize that providing model code is strongly encouraged. There are valid reasons for why researchers may not be able to provide model code, including possible intellectual property concerns (this potential consideration has been illustrated in the artificial intelligence modeling field more broadly, for example, [34]), or specific concerns about potential misuse (i.e., using the forecasting model as a “black box” without understanding its principles and limitations). In cases where code is not made available, we propose that authors provide a brief justification for why this is the case. This may assist in future studies that seek to identify and mitigate barriers to sharing forecast model code during public health emergencies. A clear statement of model code availability will also allow journals to screen submissions for this feature.

D. Model evaluation

Item 10: Describe the model validation, and justify the approach. We recommend describing the model validation method and justifying the approach. Forecast model validation is critical to ensure accuracy of results and usefulness of models, and it also encourages trust in the results and methods by other researchers, journal reviewers, journal editors, and end users. Forecasting research should indicate if cross-validation or out-of-sample validation was performed, the data used for validation, how many models were considered at each stage of validation, the time span of validation (with justification), and whether the researchers were blinded to the external validation dataset (for example, through a prospective design like a forecast challenge or other real-time forecasting exercise) [2,5,35–37].

Item 11: Describe the forecast accuracy evaluation method used, with justification. We recommend that the forecast accuracy evaluation method is described and justified. Forecast and prediction research studies may include point predictions (for example, mean number of expected cases) or a full probability distribution of the outcome of interest. It is important that the metric of validation accuracy is both clearly defined and justified, thereby allowing forecast performance to be robustly evaluated and compared between studies when using the same data. Examples include [22].

Item 12: Where possible, compare results to a benchmark or other comparator model, with justification of comparator choice. We recommend that the forecast results are compared to a benchmark or other comparator model, and the comparator model choice justified. Benchmark models may include relatively simple models such as autoregression or seasonal

averages [38]. These comparisons are important to mitigate the risk of model misspecification and may also provide a “common sense” interpretation of forecast value compared to intuitive benchmarks such as an autoregression model with a 1-week lag time [39–41]. If there are other published models for the specific forecasting target or type of target that demonstrate significant improvement compared to simpler models, those forecasts should be used as the comparator to the extent possible [20]. Comparison may include formal statistical comparisons with established methods (for example, Diebold-Mariano tests or permutation tests) [42,43]. For emerging pathogens with novel disease traits and novel forecasting targets, such benchmark models may not be readily available.

Item 13: Describe the forecast horizon, with justification of its length. We recommend that the forecast horizon is described, and its length justified (for example, [21]). Presenting forecast accuracy and precision with increasing lead times allows for an evaluation of a forecast’s usefulness over operationally relevant timescales. We suggest justification of the forecast horizon to avoid inadvertent misrepresentation of model accuracy and precision and to communicate the inherent limits of forecasts that may break down over longer forecast horizons [38,41].

Item 14: Present and explain uncertainty of forecasting results. We recommend that uncertainty in the forecasting results is presented and explained. Uncertainty is a fundamental consideration in developing and interpreting epidemic forecasting and prediction research. Uncertainty can arise from parameters, assumptions, model choice, lack of knowledge about the epidemiology of the disease, or variability in the data itself. Qualitative and/or quantitative estimates of uncertainty can be incorporated into forecasting research through using probabilistic forecast methods, uncertainty intervals around point estimates (for example, prediction or credibility intervals), sensitivity or scenario analyses, or description of the uncertainty in the model parameters. We recommend that the estimates of uncertainty are clearly described in at least the results, and, ideally, also referred to in the discussion and the abstract. Examples include [20,22].

E. Translation of results for public health practice, interpretability, and generalizability

Item 15: Briefly summarize the results in nontechnical terms, including a nontechnical interpretation of forecast uncertainty. We recommend that the results are summarized in nontechnical terms. Adequately reporting and explaining model forecasts is critical for a wide range of readers, including public health decision-makers and the media. Forecasts can be misinterpreted, especially when uncertainty is not explicitly and clearly communicated with a broad audience in mind. We propose that a lack of appropriate communication about these inherent caveats in forecasting science may lead to skepticism of forecasting by important end users (such as decision-makers), the media, and the general public. We recommend a brief nontechnical summary of forecasting research results, as already required by several major biomedical journals for a range of research fields [44], and including a nontechnical interpretation of forecast uncertainty. Examples include [45,46].

Item 16: If results are published as a data object, encourage a time-stamped version number. In general, we recommend that results, i.e., the raw forecast data themselves and summaries of these data, are made available as a public data object. This reporting recommendation serves multiple purposes. First, it allows searching and aggregating of forecast results by a standardized object nomenclature. Second, it ensures that forecasts are truly prospective, when claimed to be so. Third, it permits clear communication of when forecasts are updated (for instance, as parameter estimates are refined, or as new data becomes available). We

recommend assigning a unique and persistent identifier to the time-stamped and versioned data object, such as a digital object identifier (DOI). This practice could extend to web-based forecasting tools linked to the publication also.

Item 17: Describe the weaknesses of the forecast, including weaknesses specific to data quality and methods. We recommend that the weaknesses in the forecast be described. Limitations can include data quality (for example, heterogeneity in sampling over time and across populations, diagnostic limitations, or case selection bias), parameter uncertainty, model misspecification, or limitations in generalizability. No model is a complete representation of reality, and much can be gleaned about a forecasting model's utility from knowing its limitations or simplifying assumptions. It is important to note that identifying methodological weaknesses in forecasts does not necessarily mean that they lack credibility. Rather, highlighting such weaknesses may inform data needs, lead to improvements of forecasts, and assist in interpretation of forecast results during public health decision-making. For example, [22,46].

Item 18: If the forecast research is applicable to a specific epidemic, comment on its potential implications and impact for public health action and decision-making. When forecasting research is intended to be applicable to a specific outbreak or epidemic, we propose that the potential implications of the forecast for that specific epidemic need to be described, including whether it has a possible impact on public health action or decision-making. Framing the discussion of results in this context is essential for model end users and may assist in ensuring that model developers are addressing the right research questions from the outset. For example, [22].

Item 19: If the forecast research is applicable to a specific epidemic, comment on how generalizable it may be across populations. When forecasting research is intended to be applicable to a specific outbreak or epidemic, researchers should describe the generalizability of results between countries, regions, populations, and perhaps even pathogens, together with the rationale for why (for example, [22]). A forecast's accuracy or applicability in one setting may not translate to others due to inherent differences in healthcare capacity, population demography, disease ecology, socioeconomic factors, and data availability and reliability.

Conclusions

We present the first guidelines for standard reporting of epidemic forecasting research, comprising 19 preferred items in a checklist. We stress that the objectives of these guidelines are intended to improve the epidemic forecasting reporting consistency and reproducibility, as well as comparability and quality. They serve as a set of standards to ensure that critical aspects of these studies are adequately reported and are not intended to advise scientists on how to perform epidemic forecast and prediction research. We note that our Delphi process also led to several checklist items, which pertain to the translation of forecasting results for public health practice.

The primary target audience of these guidelines is scientists using models to forecast infectious disease epidemics as a means to ensure that critical reporting items are included in published manuscripts. While this checklist may also serve as a means of ensuring standardization of infectious disease modeling quality among this group, it is distinct from other structured consensus documents, which have focused on modeling principles or made recommendations for reporting of other types of modeling studies [47–50]. The secondary target audience of these guidelines include model users (for example, those in operational public health and policy), journal peer reviewers, journal editors, and epidemiology training programs. We encourage formal endorsement by modeling groups and broad adoption by biomedical journals who already require completion of reporting checklists for manuscript submissions, including

clinical trials and systematic reviews [51]. While our guidelines were developed with peer-reviewed published research papers in mind, these could be applied to epidemic forecasting research reported elsewhere.

Research reporting guidelines do need to be subfield specific to be pragmatic and useful (for this reason the EQUATOR website references over 440 guidelines), but it is worth comparing our final guidelines to others medical reporting guidelines, which have been widely implemented, such as CONSORT and PRISMA [14]. Like EPIFORGE, these guidelines identify the study type, define the study objectives, comment on study limitations, aid in interpreting the validity of the results, and discuss generalizability of the findings.

Our approach to development of this checklist has some limitations. While the major strength of the EPIFORGE guidelines is the use of a structured Delphi process across a range of stakeholders, this resulted in a number of valuable reporting considerations suggested by the Delphi panel, which were not included after the consensus process. We noted several items suggested by the Delphi panel that were not ultimately voted in. These covered a range of topics and may not be applicable to all forecasting and prediction research. We include these items as a supporting appendix for general consideration in the field of reporting forecasting and prediction research, and these may be reconsidered in future versions of the EPIFORGE reporting guidelines (S2 Table). Future versions should also seek to identify other items which through a new Delphi process.

While the development process involved broad consultation, we encourage broad and frank feedback and critique. Feedback will be valuable in updating future iterations of these guidelines, which are intended to be dynamic and responsive to the ongoing needs of epidemic forecasters and end users, including those involved in COVID-19 research and response. These guidelines have been submitted to the EQUATOR Network webpage, in addition to dedicated webpages to facilitate feedback and journal endorsement ([20–22]; <https://midasnetwork.us/>), following examples from other guidelines [14].

Supporting information

S1 Text. Full description of methods.

(DOCX)

S1 Table. Delphi panel members.

(DOCX)

S2 Table. Other reporting considerations not included into the final reporting checklist.

(DOCX)

Acknowledgments

We appreciate the role of the Outbreak Science and Model Implementation Working Group in developing this initiative, and the Johns Hopkins Center for Health Security for hosting the face-to-face consensus meeting and conducting the electronic Delphi process.

Disclaimers

The content is solely the responsibility of the authors and does not necessarily represent the official views of NIGMS or the National Institutes of Health.

Material has been reviewed by the Walter Reed Army Institute of Research. There is no objection to its presentation and/or publication. The opinions or assertions contained herein

are the private views of the author, and are not to be construed as official, or as reflecting true views of the Department of the Army or the Department of Defense.

The views expressed here are those of the authors and do not necessarily reflect the official policy of the Department of Defense, Department of the Army, US Army Medical Department, or the US Government.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

This material was approved by Los Alamos National Laboratory (LA-UR-20-24531), which is operated by Triad National Security, LLC, for the National Nuclear Security Administration of the US Department of Energy (Contract No. 89233218CNA000001).

References

1. Lowe R, Stewart-Ibarra AM, Petrova D, García-Díez M, Borbor-Cordova MJ, Mejía R, et al. Climate services for health: predicting the evolution of the 2016 dengue season in Machala. Ecuador *Lancet Planet Health*. 2017; 1:e142–51. [https://doi.org/10.1016/S2542-5196\(17\)30064-5](https://doi.org/10.1016/S2542-5196(17)30064-5) PMID: 29851600
2. Johansson MA, Apfeldorf KM, Dobson S, Devita J, Buczak AL, Baugher B, et al. An open challenge to advance probabilistic forecasting for dengue epidemics. *Proc Natl Acad Sci*. 2019; 116:24268–74. <https://doi.org/10.1073/pnas.1909865116> PMID: 31712420
3. Biggerstaff M, Johansson M, Alper D, Brooks LC, Chakraborty P, Farrow DC, et al. Results from the Second Year of a Collaborative Effort to Forecast Influenza Seasons in the United States. *Epidemics*. 2018; 24:26–33. <https://doi.org/10.1016/j.epidem.2018.02.003> PMID: 29506911
4. Ahrens KA, Hutcheon JA, Gavin L, Moskosky S. Reducing Unintended Pregnancies as a Strategy to Avert Zika-Related Microcephaly Births in the United States: A Simulation Study. *Matern Child Health J*. 2017; 21:982–7. <https://doi.org/10.1007/s10995-017-2275-2> PMID: 28102503
5. Del Valle SY, McMahon BH, Asher J, Hatchett R, Lega JC, Brown HE, et al. Summary results of the 2014–2015 DARPA Chikungunya challenge. *BMC Infect Dis*. 2018; 18:245. <https://doi.org/10.1186/s12879-018-3124-7> PMID: 29843621
6. Evans MV, Dallas TA, Han BA, Murdock CC, Drake JM. Data-driven identification of potential Zika virus vectors. Brady O, editor. *eLife*. 2017; 6: e22053. <https://doi.org/10.7554/eLife.22053> PMID: 28244371
7. Funk S, Camacho A, Kucharski AJ, Eggo RM, Edmunds WJ. Real-time forecasting of infectious disease dynamics with a stochastic semi-mechanistic model. *Epidemics*. 2018; 22:56–61. <https://doi.org/10.1016/j.epidem.2016.11.003> PMID: 28038870
8. Kobres P-Y, Chretien J-P, Johansson MA, Morgan JJ, Whung P-Y, Mukundan H, et al. A systematic review and evaluation of Zika virus forecasting and prediction research during a public health emergency of international concern. Pimenta PFP, editor. *PLoS Negl Trop Dis*. 2019; 13:e0007451. <https://doi.org/10.1371/journal.pntd.0007451> PMID: 31584946
9. Rainisch G, Asher J, George D, Clay M, Smith TL, Kosmos C, et al. Estimating Ebola Treatment Needs, United States. *Emerg Infect Dis*. 2015; 21:1273–5. <https://doi.org/10.3201/eid2107.150286> PMID: 26079150
10. Moghadas SM, Shoukat A, Fitzpatrick MC, Wells CR, Sah P, Pandey A, et al. Projecting hospital utilization during the COVID-19 outbreaks in the United States. *Proc Natl Acad Sci*. 2020; 117:9122–6. <https://doi.org/10.1073/pnas.2004064117> PMID: 32245814
11. Alwan NA, Bhopal R, Burgess RA, Colburn T, Cuevas LE, Smith GD, et al. Evidence informing the UK's COVID-19 public health response must be transparent. *Lancet*. 2020; 395:1036–7. [https://doi.org/10.1016/S0140-6736\(20\)30667-X](https://doi.org/10.1016/S0140-6736(20)30667-X) PMID: 32197104
12. Popper KR, Karl RS. *Logic of scientific discovery*. London: Hutchinson and Co; 1959.
13. Browse reporting guidelines by specialty | The EQUATOR Network. [cited 2021 Aug 13]. Available from: <https://www.equator-network.org/library/browse-reporting-guidelines-by-specialty/>.
14. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021; n71. <https://doi.org/10.1136/bmj.n71> PMID: 33782057
15. Hopewell S, Dutton S, Yu L-M, Chan A-W, Altman DG. The quality of reports of randomised trials in 2000 and 2006: comparative study of articles indexed in PubMed. *BMJ*. 2010; 340:697–7. <https://doi.org/10.1136/bmj.c723> PMID: 20332510

16. Behrend MR, Basáñez M-G, Hamley JID, Porco TC, Stolk WA, Walker M, et al. Modelling for policy: The five principles of the Neglected Tropical Diseases Modelling Consortium. *PLoS Negl Trop Dis*. 2020; 14:e0008033. <https://doi.org/10.1371/journal.pntd.0008033> PMID: 32271755
17. Pollett S, Johansson M, Biggerstaff M, Morton LC, Bazaco SL, Brett Major DM, et al. Identification and evaluation of epidemic prediction and forecasting reporting guidelines: A systematic review and a call for action. *Epidemics*. 2020; 33:100400. <https://doi.org/10.1016/j.epidem.2020.100400> PMID: 33130412
18. Moher D, Schulz KF, Simera I, Altman DG. Guidance for Developers of Health Research Reporting Guidelines. *PLoS Med*. 2010; 7:e1000217. <https://doi.org/10.1371/journal.pmed.1000217> PMID: 20169112
19. The EQUATOR Network | Enhancing the QUALity and Transparency Of Health Research. [cited 2021 Aug 13]. Available from: <https://www.equator-network.org/>.
20. Osthus D, Gattiker J, Priedhorsky R, Valle SYD. Dynamic Bayesian Influenza Forecasting in the United States with Hierarchical Discrepancy (with Discussion). *Bayesian Anal*. 2019; 14:261–312. <https://doi.org/10.1214/18-BA1117>
21. Generous N, Fairchild G, Deshpande A, Valle SYD, Priedhorsky R. Global Disease Monitoring and Forecasting with Wikipedia. *PLoS Comput Biol*. 2014; 10:e1003892. <https://doi.org/10.1371/journal.pcbi.1003892> PMID: 25392913
22. Colón-González FJ, Soares Bastos L, Hofmann B, Hopkin A, Harpham Q, Crocker T, et al. Probabilistic seasonal dengue forecasting in Vietnam: A modelling study using superensembles. *PLoS Med*. 2021; 18:e1003542. <https://doi.org/10.1371/journal.pmed.1003542> PMID: 33661904
23. Massad E, Nascimento Burattini M, Khan K, Struchiner CJ, Coutinho FAB, Wilder-Smith A. On the origin and timing of Zika virus introduction in Brazil. *Epidemiol Infect*. 2017; 145:2303–12. <https://doi.org/10.1017/S0950268817001200> PMID: 28675351
24. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*. 2020; 368:860–8. <https://doi.org/10.1126/science.abb5793> PMID: 32291278
25. McGowan CJ, Biggerstaff M, Johansson M, Apfeldorf KM, Ben-Nun M, Brooks L, et al. Collaborative efforts to forecast seasonal influenza in the United States, 2015–2016. *Sci Rep*. 2019; 9:683. <https://doi.org/10.1038/s41598-018-36361-9> PMID: 30679458
26. Reich NG, McGowan CJ, Yamana TK, Tushar A, Ray EL, Osthus D, et al. Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the U.S. *PLoS Comput Biol*. 2019; 15:e1007486. <https://doi.org/10.1371/journal.pcbi.1007486> PMID: 31756193
27. Rainisch G, Shankar M, Wellman M, Merlin T, Meltzer MI. Regional Spread of Ebola Virus, West Africa, 2014. *Emerg Infect Dis*. 2015; 21:444–7. <https://doi.org/10.3201/eid2103.141845> PMID: 25693782
28. Chretien JP, Rivers CM, Johansson MA. Make Data Sharing Routine to Prepare for Public Health Emergencies. *PLoS Med*. 2016; 13:e1002109. <https://doi.org/10.1371/journal.pmed.1002109> PMID: 27529422
29. WHO's code of conduct for open and timely sharing of pathogen genetic sequence data during outbreaks of infectious disease (draft). Available from: https://www.who.int/blueprint/what/norms-standards/GSDDraftCodeConduct_forpublicconsultation-v1.pdf.
30. Rivers C, Chretien J-P, Riley S, Pavlin JA, Woodward A, Brett-Major D, et al. Using “outbreak science” to strengthen the use of models during epidemics. *Nat Commun*. 2019; 10:1–3. <https://doi.org/10.1038/s41467-018-07882-8> PMID: 30602773
31. Data Availability. *PLoS ONE* [Internet]. Available from: <https://journals.plos.org/plosone/s/data-availability>.
32. Porgo TV, Norris SL, Salanti G, Johnson LF, Simpson JA, Low N, et al. The use of mathematical modeling studies for evidence synthesis and guideline development: A glossary. *Res Synth Methods*. 2019; 10:125–33. <https://doi.org/10.1002/jrsm.1333> PMID: 30508309
33. Perkins TA, Siraj AS, Ruktanonchai CW, Kraemer MUG, Tatem AJ. Model-based projections of Zika virus infections in childbearing women in the Americas. *Nat Microbiol*. 2016; 1. <https://doi.org/10.1038/nmicrobiol.2016.126> PMID: 27562260
34. Frequently Asked Questions: AI and IP Policy. [cited 2021 Aug 13]. Available from: https://www.wipo.int/about-ip/en/artificial_intelligence/faq.html.
35. Venkatramanan S, Lewis B, Chen J, Higdon D, Vullikanti A, Marathe M. Using data-driven agent-based models for forecasting emerging infectious diseases. *Epidemics*. 2018; 22:43–9. <https://doi.org/10.1016/j.epidem.2017.02.010> PMID: 28256420

36. Lowe R, Barcellos C, Coelho CAS, Bailey TC, Coelho GE, Graham R, et al. Dengue outlook for the World Cup in Brazil: an early warning model framework driven by real-time seasonal climate forecasts. *Lancet Infect Dis*. 2014; 14:619–26. [https://doi.org/10.1016/S1473-3099\(14\)70781-9](https://doi.org/10.1016/S1473-3099(14)70781-9) PMID: 24841859
37. Lowe R, Coelho CA, Barcellos C, Carvalho MS, Catão RDC, Coelho GE, et al. Evaluating probabilistic dengue risk forecasts from a prototype early warning system for Brazil. *eLife*. 5:e11285. <https://doi.org/10.7554/eLife.11285> PMID: 26910315
38. Reich NG, Brooks LC, Fox SJ, Kandula S, McGowan CJ, Moore E, et al. A collaborative multiyear, multimodel assessment of seasonal influenza forecasting in the United States. *Proc Natl Acad Sci*. 2019; 116:3146. <https://doi.org/10.1073/pnas.1812594116> PMID: 30647115
39. Pollett S, Boscardin WJ, Azziz-Baumgartner E, Tinoco YO, Soto G, Romero C, et al. Evaluating Google Flu Trends in Latin America: Important Lessons for the Next Phase of Digital Disease Detection. *Clin Infect Dis*. 2017; 64:34–41. <https://doi.org/10.1093/cid/ciw657> PMID: 27678084
40. Olson DR, Konty KJ, Paladini M, Viboud C, Simonsen L. Reassessing Google Flu Trends data for detection of seasonal and pandemic influenza: a comparative epidemiological study at three geographic scales. *PLoS Comput Biol*. 2013; 9:e1003256. <https://doi.org/10.1371/journal.pcbi.1003256> PMID: 24146603
41. McGough SF, Brownstein JS, Hawkins JB, Santillana M. Forecasting Zika Incidence in the 2016 Latin America Outbreak Combining Traditional Disease Surveillance with Search, Social Media, and News Report Data. *PLoS Negl Trop Dis* 2017; 11. <https://doi.org/10.1371/journal.pntd.0005295> PMID: 28085877
42. Diebold FX, Mariano RS. Comparing Predictive Accuracy. *J Bus Econ Stat*. 1995; 13:253. <https://doi.org/10.2307/1392185>
43. Ray EL, Reich NG. Prediction of infectious disease epidemics via weighted density ensembles. *PLoS Comput Biol*. 2018; 14:e1005910. <https://doi.org/10.1371/journal.pcbi.1005910> PMID: 29462167
44. Shailes S. Something for everyone. *eLife*. 2017; 6:e25411. <https://doi.org/10.7554/eLife.25411> PMID: 28294940
45. Goldstein E, Cobey S, Takahashi S, Miller JC, Lipsitch M. Predicting the Epidemic Sizes of Influenza A/H1N1, A/H3N2, and B: A Statistical Method. *PLoS Med*. 2011; 8:e1001051. <https://doi.org/10.1371/journal.pmed.1001051> PMID: 21750666
46. Lowe R, Gasparrini A, Meerbeeck CJV, Lippi CA, Mahon R, Trotman AR, et al. Nonlinear and delayed impacts of climate on dengue risk in Barbados: A modelling study. *PLoS Med*. 2018; 15:e1002613. <https://doi.org/10.1371/journal.pmed.1002613> PMID: 30016319
47. Caro JJ, Briggs AH, Siebert U, Kuntz KM, ISPOR-SMDM Modeling Good Research Practices Task Force. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. *Med Decis Mak*. 2012; 32:667–677. <https://doi.org/10.1177/0272989X12454577> PMID: 22990082
48. Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, Paltiel AD, et al. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. *Med Decis Mak*. 2012; 32:722–32. <https://doi.org/10.1177/0272989X12458348> PMID: 22990087
49. Dahabreh IJ, Trikalinos TA, Balk EM, Wong JB. Guidance for the Conduct and Reporting of Modeling and Simulation Studies in the Context of Health Technology Assessment. *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK396066/>.
50. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force—5. *Value Health J Int Soc Pharmacoeconomics Outcomes Res*. 2012; 15:828–34. <https://doi.org/10.1016/j.jval.2012.06.011> PMID: 22999132
51. Submission Guidelines. *PLoS Med* [Internet]. Available from: <https://journals.plos.org/plosmedicine/s/submission-guidelines#loc-guidelines-for-specific-study-types>.