Caution Warranted: Using the Institute for Health Metrics and Evaluation Model for Predicting the Course of the COVID-19 Pandemic

Nicholas P. Jewell, PhD; Joseph A. Lewnard, PhD; and Britta L. Jewell, PhD

A recent modeling analysis by the Institute for Health Metrics and Evaluation (IHME) (1) projecting deaths due to coronavirus disease 2019 (COVID-19) has attracted considerable attention, including from the U.S. government (2). The model used COVID-19 mortality projections to estimate hospital bed requirements and deaths. We agree with qualitative conclusions that demand for hospital beds may exceed capacity and efforts to enhance mitigation policies and surge planning are essential. Data endorse shelter-in-place orders and suggest that these measures must remain while awaiting advances in surveillance, treatment, and vaccines.

The IHME projections are based not on transmission dynamics but on a statistical model with no epidemiologic basis. Specifically, the model used reported worldwide COVID-19 deaths and extrapolated similar patterns in mortality growth curves to forecast expected deaths. The technique uses mortality data, which are generally more reliable than testing-dependent confirmed case counts. Outputs suggest precise estimates (albeit with uncertainty bounds) for all regions until the epidemic ends. This appearance of certainty is seductive when the world is desperate to know what lies ahead. However, the underlying data and statistical model must be interpreted cautiously. Here, we raise concerns about the validity and usefulness of the projections for policymakers.

First, the statistical model assumes that systematic variation in mortality curves across regions is captured by timing of social distancing decisions and that other differences are explained by random effects. The model rests on the likely incorrect assumption that effects of social distancing policies are the same everywhere and that suppression policies will be implemented in all regions and will remain effective throughout.

These projections may suffer from the fallacy of Farr's law, a similar nonmechanistic method in which epidemics are assumed to follow a normal distribution shifted and scaled to fit data. However, epidemic curves fit early data in multiple ways that affect expected durations or maximum numbers of deaths observed (3, 4). This is important as epidemics progress and deviations from normal distributions are expected–for example, due to "second waves" after interventions are eased.

Second, the approach models mortality curves for every region with parameters for the final total, the pace of mortality growth, and timing of when the growth curve inflects. We have observed few entire curves; Hubei province's is the most complete, and curves in regions of Italy and in South Korea have subsequently passed their peaks. After age and subtle differences in policy timings are accounted for, all curves

are assumed to follow these general patterns. This is optimistic: China enacted more stringent restrictions than elsewhere after observing only 17 deaths (5), and South Korea benefited from widespread testing to isolate cases early. Updating results may diminish the extent to which inference depends on a few settings, but countries that have flattened death curves earliest may not provide a basis for extrapolating trends in areas where similar control could prove elusive. Moreover, recrudescence of transmission remains possible between transient intervention periods.

Third, death counts can be unreliable and reporting differences occur even within regions. Although Hubei province data represent the most complete mortality curve available, these numbers are suspect (6). Italy and other countries only report hospital deaths, thus neglecting deaths elsewhere (7). Lack of testing may impede deaths from being attributed to COVID-19, particularly initially. For example, in Bergamo, Italy, the number of anomalous deaths is several times larger than the official COVID-19 numbers (8). Reporting delays understate the growth of mortality curves initially, which is particularly concerning because the model uses early patterns for future projections. The recent addition of aggregate hospitalizations is also problematic due to inconsistent and poor reporting.

Although undercounting of deaths affects the final epidemic size, it may not affect *shapes* of mortality curves. However, this assumes that undercounting and reporting delays are similar across time and geography, with variations captured by posited random effects. Such reporting issues resemble those arising commonly in infectious disease surveillance analyses and should be accounted for statistically.

Fourth, despite not accounting for data and model structure flaws, uncertainty bands are broad. If all sources of uncertainty were accommodated, confidence intervals would necessarily be wider, making projections less proscriptive for policy decisions. Unaccounted sources of uncertainty arise from inaccurate temporal data on mortality and hospitalization counts; model misspecification, including parametrization choices; and inaccuracies in assumptions regarding the timing and effect of social distancing policies across regions. Graphical representation of uncertainty in curves is also not conducive to understanding uncertainty in peak daily death or hospital admission dates. This uncertainty would be more evident if only the "envelope" of uncertainty was shown without the central curve, which currently suggests greater precision than the model is able to offer.

Fifth, updated projections already reveal substantial volatility. For New York, the model predicted 10 243

deaths (range, 5167 to 26 444) on 27 March and 15 546 (range, 8016 to 22 255) on 30 March. Given the opaqueness of the model and underlying source data, it is challenging to understand why other regions' projections also change dramatically. The alignment of past predictions with reality and current predictions should also be reported transparently.

Finally, the projections are being interpreted misleadingly in formal and social media, without sufficient caveats, and outcomes differ substantially from those of other models (9, 10). Upper uncertainty bounds are being interpreted as "worst-case" when, at best, they reflect only one scenario. The model yields an attack rate less than 5%; higher rates will lead to greater mortality than upper bounds.

Ultimately, IHME's model may be reliable only for short-term projections. For hospital demand projections, patient-level clinical outcome data will enable more accurate conclusions than poorly reported worldwide aggregate mortality data with point estimates of how deaths translate into hospital use. Local data are less likely to be subject to undercounting or reporting errors, helping hospitals better prepare for the immediate future. It is also unlikely that a "one-size" model will fit all regions at all times. Policymakers will be best served when they consider projections from multiple models, thus increasing the understanding of factors that influence disparate projections and enhancing comprehension of unaccounted uncertainty in any one model. Major policy decisions need model input, but models are valuable only to the extent that outputs are transparent, are valid, are based on accurate documented sources, are rigorously evaluated, and yield robust and reliable projections.

From London School of Hygiene & Tropical Medicine, London, United Kingdom, and University of California, Berkeley, Berkeley, California (N.P.J., J.A.L.); University of California, Berkeley, Berkeley, California; and Imperial College London, London, United Kingdom (B.L.J.).

Corresponding Author: Nicholas P. Jewell, PhD, Department of Medical Statistics, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom; e-mail, jewell@berkeley.edu.

Current author addresses and author contributions are available at Annals.org.

References

1. **Institute for Health Metrics and Evaluation.** New COVID-19 forecasts: US hospitals could be overwhelmed in the second week of April by demand for ICU beds, and US deaths could total 81,000 by July. Accessed at www.healthdata.org/news-release/new-covid-19 -forecasts-us-hospitals-could-be-overwhelmed-second-week-april -demand-icu on 30 March 2020.

2. **Institute for Health Metrics and Evaluation (IHME).** IHME | The Chris Murray Model | Dr. Deborah Birx | COVID-19 Model | "Meet The Press". 31 March 2020. Accessed at www.youtube.com/watch? v=vnd_wITL8_4 on 3 April 2020.

3. **Bregman DJ, Langmuir AD.** Farr's law applied to AIDS projections. JAMA. 1990;263:1522-5. [PMID: 2308183] doi:10.1001/jama.1990 .03440110088033

4. **@CT_Bergstrom.** 1. A somewhat more technical thread than usual. 28 March 2020. Accessed at https://twitter.com/CT_Bergstrom /status/1243819232950751233 on 13 April 2020.

5. **Kuo L.** Coronavirus: China bars 11m residents from leaving city at centre of outbreak. The Guardian. 22 January 2020. Accessed at www .theguardian.com/world/2020/jan/22/coronavirus-china-measures-rein -spread-mutate-disease-death-toll on 31 March 2020.

6. **Long Q, Siu-fung L, Mudie L.** Estimates Show Wuhan Death Toll Far Higher Than Official Figure. Radio Free Asia. 27 March 2020. Accessed at www.rfa.org/english/news/china/wuhan -deaths-03272020182846.html on 31 March 2020.

7. **Stancati M, Sylvers E.** Italy's Coronavirus Death Toll Is Far Higher Than Reported. The Wall Street Journal. 1 April 2020. Accessed at www.wsj.com/articles/italys-coronavirus-death-toll-is-far-higher-than -reported-11585767179 on 4 April 2020.

8. **Foresti CCL.** The real death toll for Covid-19 is at least 4 times the official numbers. Corriere Della Serra. 26 March 2020. Accessed at www.corriere.it/politica/20_marzo_26/the-real-death-toll-for-covid -19-is-at-least-4-times-the-official-numbers-b5af0edc-6eeb-11ea -925b-a0c3cdbe1130.shtml?refresh_ce-cp on 31 March 2020.

9. **Ferguson NM, Laydon D, Nedjati-Gilani G, et al.** Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College London; 16 March 2020. Accessed at www.imperial.ac.uk/media/imperial -college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf on 31 March 2020.

10. **Kerr C, Hagedorn B, Mistry D, et al.** Working paper: Projected COVID-19 epidemic trends and health system needs for Oregon. Institute for Disease Modeling; 23 March 2020. Accessed at https ://covid.idmod.org/data/Projected_COVID19_epidemic_trends _health_system_needs_Oregon.pdf on 31 March 2020.

Current Author Addresses: Dr. N.P. Jewell: Department of Medical Statistics, London School of Hygiene & Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom.
Dr. Lewnard: Division of Epidemiology and Biostatistics, School of Public Health, University of California, Berkeley, Room 5410, 2121 Berkeley Way West, Berkeley, CA 94720.
Dr. B.L. Jewell: Department of Infectious Disease Epidemiology, Imperial College London, Norfolk Place, London W2 1PG, United Kingdom.

Author Contributions: Conception and design: N.P. Jewell, J.A. Lewnard, B.L. Jewell.
Analysis and interpretation of the data: B.L. Jewell.
Drafting of the article: N.P. Jewell, B.L. Jewell.
Critical revision of the article for important intellectual content: N.P. Jewell, J.A. Lewnard, B.L. Jewell.
Final approval of the article: N.P. Jewell, J.A. Lewnard, B.L. Jewell.
Statistical expertise: N.P. Jewell.
Collection and assembly of data: B.L. Jewell.