Disaster Medicine and Public Health Preparedness

www.cambridge.org/dmp

Letter to the Editor

Cite this article: Collins GS (2022) COVID-19 prediction models need robust and transparent development. *Disaster Med Public Health Prep* **16**: 1318–1319. doi: https://doi.org/10.1017/dmp.2021.119.

First published online: 19 April 2021

Keywords:
prediction; covid-19; pandemics

Corresponding author:
Gary S. Collins,
Email: gary.collins@csm.ox.ac.uk.

COVID-19 Prediction Models Need Robust and Transparent Development

Gary S. Collins PhD

Centre for Statistics in Medicine, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK

The 2019 coronavirus disease (COVID-19) pandemic has led researchers to develop prediction models for COVID-19 to aid diagnosis and prognosis. To date, 145 COVID-19 prediction models have been developed, and all have been critically rated at high risk of bias with flaws in the design, statistical analysis, and reporting.[1] The recent study by Niu and colleagues describes the development of a new prediction model to identify COVID-19 patients at increased risk of dying.[2] Disappointingly, this study shares many common features of the aforementioned existing models.

Sample size is a key design feature to ensure that sufficient numbers of participants are included to meet the study objectives. The study by Niu and colleagues did not include a sample size calculation. The study used published sample size formulas for developing prediction models,[3] based on information reported in Niu et al.'s study, namely, examining 69 predictors, an outcome prevalence of 31/150 = 0.21, along with a conservative (unreported) estimate of the anticipated model R^2 (based on 15% of the maximum possible R^2, which is dependent on the outcome prevalence, ie, 15% of 0.64 = 0.096), the minimum sample size in 6118 individuals (1265 events). The study by Niu and colleagues included 150 individuals, where 31 individuals died, a substantially lower than the required sample size. To precisely estimate the intercept alone requires 252 individuals (52 events), again larger than the sample size included in their study.

The consequences of an inadequate sample size to develop a prediction model are the risk of overfitting, such that the model fits idiosyncrasies in the development data, yet fails to work in new data. This is illustrated by their near perfect area under the curve (AUC) of 0.97, a value that is undoubtedly overestimated. An internal validation using bootstrapping (not carried in this study, though recommended),[4] would have shown the actual AUC to be substantially lower. Although Niu and colleagues evaluated their model in an external data set, the sample size of these data was insufficient. Sample size considerations suggest a minimum of 100 outcome events; Niu et al.'s study included an external data set containing only 12 deaths, making the drawing unreliable of any conclusions on predictive accuracy.

The next concern relates to the handling of continuous variables, where all variables were dichotomized – a practice widely discredited.[4] Dichotomizing continuous variables is biologically implausible, with individuals on either side of the cut-point with similar values being assigned different levels of risk, yet individuals at the lower and upper end in each group with different values were assigned with the same risk. Furthermore, dichotomizing continuous measurements discards important predictive information.

Other major concerns include the use of univariate screening, forward selection, and the handling of missing data. These issues impact on model reliability and accuracy and, along with the other issues highlighted, will lead to a model producing inaccurate predictions,[5] making the model potentially harmful to use.

Both the conduct and reporting of this study and others developing new COVID-19 prediction models would be improved with adherence to the TRIPOD Statement (www.tripod-statement.org) that outlines important information to report, as well as provides guidance on best practice to develop clinical prediction models.[4]

Conflict(s) of Interest. The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this paper.

© Society for Disaster Medicine and Public Health, Inc. 2021.

References

1. **Wynants L, Van Calster B, Collins GS, *et al.*** Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. *BMJ.* 2020;369:m1328. doi: 10.1136/bmj.m1328.
2. **Niu Y, Zhan Z, Li J, *et al.*** Development of a predictive model for mortality in hospitalized patients with COVID-19. Disaster Med Public Health Prep. 2021;epub:1-33. doi: 10.1017/dmp.2021.8.
3. **Riley RD, Ensor J, Snell KIE, *et al.*** Calculating the sample size required for developing a clinical prediction model. *BMJ.* 2020;368:m441. doi: 10.1136/bmj.m441.

4. **Moons KGM, Altman DG, Reitsma JB, *et al.*** Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1): W1–W73.

5. **Steyerberg EW, Uno H, Ioannidis JPA, van Calster B**; and Collaborators. Poor performance of clinical prediction models: the harm of commonly applied methods. *J Clin Epidemiol.* 2018;98:133–143. doi: 10.1016/j.jclinepi.2017.11.013.

https://doi.org/10.1017/dmp.2021.119 Published online by Cambridge University Press