Applications of predictive modelling early in the COVID-19 epidemic

On Jan 30, 2020, WHO declared a Public Health Emergency of International Concern, a month after COVID-19 was identified in Wuhan, China. By this point, several mathematical and computational models had already raised the alarm about the potential for the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) to cause a global pandemic and the dire consequences for public health should drastic action not be taken. During the emergence of a novel pandemic, predictive modelling is important in public health planning and response. Relating models to data provides a view into unseen variables, such as the occurrence of cryptic transmission and the prevalence of infection, and these models allow exploration of counterfactuals and hypothetical interventions. However, although there have been tremendous advances in mathematical epidemiology, prognostications about epidemic outcomes are inherently prone to errors. Predictive modelling is valuable when assumptions are related, the variables to be estimated are clearly defined, and researchers or policy makers who use the model outputs have a clear understanding of what can and cannot be achieved by this method. Indeed, calls for national disease-forecasting centres have arisen from the crucial need to educate policy makers at all levels on how to integrate predictive modelling into decision-making processes.

Deriving insights with predictive modelling requires diverse datasets, which are often imperfect, particularly in the crucial period of epidemic emergence when surveillance is imprecise and little is known about the epidemiology or the clinical features of the disease. For example, extensive clinical case counts and genomic data were combined with large-scale records of human mobility and behaviour using predictive modelling, owing in part to the massive deployment of digital information sources. In this Comment, we highlight several important discoveries resulting from the application of predictive modelling to diverse data sources that affected clinical and policy decisions.

In the weeks following the first report of COVID-19, predictive models anticipated the pattern of international spread but also quantified the extent of the epidemic in China. Specifically, a predictive model by Imai and colleagues[1] used travel volumes from Wuhan and the dates when imported cases first arrived in cities within China and globally to forecast the size of the epidemic in Wuhan. The results of this study suggested that substantially more cases were present in Wuhan than were reported in the official statistics.[1] Identifying the potential discrepancy between reported cases and true disease burden provided a crucial early warning to the international community. Next, statistical modelling and data-driven computer simulations provided accurate projections of global epidemic dispersal, quantifying the role of physical distancing in China and reductions in international travel on the spatiotemporal pattern of spread of COVID-19.[2,3] These predictive models showed that the cordon sanitaire around Wuhan reduced the growth rate of exported cases but came too late to prevent national and international seeding. Control of the epidemic in countries outside China failed because of the difficulty in detecting and isolating infected travellers. Mechanistic modelling of the natural history and transmission of COVID-19 anticipated this difficulty.[4] A predictive model provided the first evidence for the hypothesis, now widely accepted, that presymptomatic and asymptomatic infected individuals fuel local epidemics. Consequently, the majority of imported cases went undetected, generating extensive chains of local transmission.[3] Owing to the difficulties of syndromic surveillance and incomplete testing, COVID-19 mortality has often been the most easily measured, widely available, and easily compared metric for epidemic progression. Estimates of infection fatality rates generated by early studies of expatriated travellers paved the way for later efforts to characterise unknown epidemic burden using various modelling approaches that relate mortality to unknown epidemic prevalence.[5]

The unprecedented scale of non-pharmaceutical measures implemented in China and later in many countries around the world resulted in a strong variation in human behaviour. Lockdowns and physical distancing measures profoundly altered human mobility and encounters. Measuring changes in human mobility under these restrictions was essential to quantify the effect

Published Online
August 10, 2020
https://doi.org/10.1016/S2589-7500(20)30196-5

See Correspondence page e503

See Online/Comment
Lancet Glob Health 2020; published online Aug 10. https://doi.org/10.1016/S2214-109X(20)30361-2 and https://doi.org/10.1016/S2214-109X(20)30362-4

See Online/Comment
Lancet Public Health 2020; published online Aug 10. https://doi.org/10.1016/S2468-2667(20)30187-0

of public health measures on the amount of human contact and geographical extent of travel. Aggregated data from mobile phone and internet service records provided an accurate and near real-time information source. By leveraging these data, predictive modelling allowed for the assessment of mobility restrictions on the propagation of the epidemic and showed how control measures implemented in China substantially mitigated the spread of COVID-19.[6,7] As the pandemic progressed and lockdowns were implemented in many countries, analyses based on mobile phone records provided essential support to public health assessments across the different stages of lockdown implementation and release.[8]

According to Google Scholar, there have been well over 30 000 academic publications with COVID-19 in the title. Of these 30 000 papers, less than 2% indicate from the title that they use predictive modelling. Nevertheless, nearly every business, hospital, city, state, and national government has been provided with COVID-19 forecasts. This disconnect between the small but rapidly growing science around outbreak forecasting and its now widespread application creates a complex situation for researchers, clinicians, and policy makers. As a result, we echo calls for disease-forecasting centres at the national level that provide not only predictive models but also expert guidance to policy makers and the public around the interpretation of the models. We conclude that predictive modelling is not a monolithic framework nor a single methodology but rather encompasses a wide variety of statistical and mathematical models applied to diverse data to address different inference and prediction goals. How can we assess the performance of predictive modelling in guiding the global response to COVID-19? Regarding the most important application of these models, there has been notable success: predictive modelling correctly predicted that a global pandemic was probable and that there would be severe consequences for human health in the absence of strong public health measures to restrict human contact.[9]

We thank the G20 Riyadh Global Digital Health Summit hosted in Saudi Arabia to leverage the role of digital health in the fight against current and future pandemics. All authors contributed equally to the literature review, interpretation, and writing of the manuscript. We declare no competing interests.

*Chiara Poletto, Samuel V Scarpino, *Erik M Volz*
e.volz@imperial.ac.uk

Pierre Louis Institute of Epidemiology and Public Health, INSERM, Paris, France (CP); Network Science Institute, Northeastern University, Boston, MA, USA (SVS); and MRC Centre for Global Infectious Disease Analysis and Department of Infectious Disease Epidemiology, Imperial College London, London W2 1PG, UK (EMV)

1 Imai N, Dorigatti I, Cori A, Donnelly C, Riley S, Ferguson NM. Report 2—Estimating the potential total number of novel Coronavirus (2019-nCoV) cases in Wuhan City, China. Jan 22, 2020. https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-2-update-case-estimates-covid-19/ (accessed Aug 7, 2020).
2 Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 2020; **368:** 395–400.
3 Pinotti F, Di Domenico L, Ortega E, et al. Tracing and analysis of 288 early SARS-CoV-2 infections outside China: a modeling study. *PLoS Med* 2020; **17:** e1003193.
4 Hellewell J, Abbott S, Gimma A, et al. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. *Lancet Glob Health* 2020; **8:** e488–96.
5 Flaxman S, Mishra S, Gandy A, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* 2020; published online June 8. https://doi.org/10.1038/s41586-020-2405-7.
6 Buckee CO, Balsari S, Chan J, et al. Aggregated mobility data could help fight COVID-19. *Science* 2020; **368:** 145–46.
7 Kraemer MUG, Yang C-H, Gutierrez B, et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 2020; **368:** 493–97.
8 Oliver N, Lepri B, Sterly H, et al. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. *Sci Adv* 2020; **6:** eabc0764.
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. March 16, 2020. https://www.imperial.ac.uk/media/imperial-college/medicine/mrc-gida/2020-03-16-COVID19-Report-9.pdf (accessed Aug 7, 2020).