VIRAL INFECTION

Nowcasting the spread of SARS-CoV-2

A statistical framework that integrates data from a fine-scale targeted testing scheme and regular randomized surveillance surveys provides unbiased and fine-grained estimates of key SARS-CoV-2 epidemiological parameters that are critical for real-time policy decision-making.

Hagai Rossman and Eran Segal

A major challenge of the SARS-CoV-2 pandemic is the ability to obtain accurate data on the status of the epidemic for policy-makers at both the local and national levels. Awareness of the situation is crucial as policy-makers have to act quickly in a fast-moving crisis and make high-stakes decisions, including enacting non-pharmaceutical interventions (NPIs) such as mask wearing, rapid vaccination campaigns and local restrictions to control viral spread. However, obtaining an accurate epidemiological picture is challenging because most national testing schemes actively select the population to be tested (targeted testing), which introduces bias, as the tested population does not typically represent the general population. In a targeted testing scheme, individuals who are tested are usually at higher risk of being infected, as the selection procedure entails experiencing one or more symptoms, contact with an infected individual or working in high-risk environments, such as health-care facilities. Selection bias, confounders, diagnostic errors due to the imperfect sensitivity and specificity of tests, and temporal and geographic variation in testing capacity can all affect epidemiological estimates derived from this type of data¹ and may paint a very different picture that in turn influences important policy decisions. Now, reporting in this issue of *Nature Microbiology*, Nicholson et al.² propose a causal framework that models large-scale and fine-grained targeted testing data along with smaller-scale but unbiased randomized testing data (Fig. 1).

Nowcasting is an attempt to understand the current spatiotemporal state of an epidemic by including epidemiological information such as prevalence, the (R_t) and the spread of variants of concern (VOC)³. Correctly estimating infections is at the heart of epidemic nowcasting, as it contains important information about the immunity and protection of the population and the magnitude and risk of viral spread. The general public may be more familiar with statistical frameworks

Fig. 1 | Nowcasting the spread of SARS-CoV-2 by combining randomized testing surveillance data and finer-scale spatiotemporal targeted testing data. A flexible statistical framework combining randomized testing surveillance data and finer-scale spatiotemporal targeted testing data provides debiased local epidemiological parameters of interest, such as estimates of SARS-CoV-2 prevalence and the effective reproduction number (R_t). The estimates from this framework provide valuable information to nowcast the epidemiological situation and provide policy makers with tools to make high-stakes decisions during viral spread. The framework addresses ascertainment bias and variations in testing capacity, allows the input of different types of tests with different characteristics (such as lateral flow devices in place of PCR tests), and generally allows customization and modular combination with other models. This work highlights that randomized testing can be an important tool in epidemic control and that it has been a largely undervalued and underused public health resource in the pandemic so far.

for nowcasting from the weather domain, in which various models have been continuously improved over the years to provide accurate weather information on different spatiotemporal levels.

In the context of SARS-CoV-2, testing at scale, that is, at the population level and at high frequency, has been paramount in epidemic nowcasting and has been applied in different settings from targeted contact tracing to population-scale testing schemes⁴.

A unique and valuable population testing scheme was rolled out in the UK early in the pandemic, comprising two randomized surveys: the Office for National Statistics COVID-19 Infection Survey (ONS CIS) and the REal-time Assessment of Community Transmission (REACT) studies, both of which aimed to test a representative sample of the population over time.

Nicholson et al. propose a bespoke Bayesian modelling framework to combine

all of the information from different testing data sources to achieve unbiased estimates of key epidemiological parameters — the prevalence and the time-varying R_t . Of note, the framework addresses selection bias and variations in testing capacity; allows for the input of different types of tests with different characteristics (such as lateral flow devices in place of PCR tests); and generally allows modular combination and customization with other models — in the article, the authors demonstrated how to incorporate an epidemiological model that accounts for vaccine-induced immunity in the population.

The modelling approach was validated and produced data that were similar to gold-standard data, including similar prevalence estimates and predicted R_t , and that were in accordance with external estimates from different models generated by other modelling efforts. The model also showed that the local R_t estimate strongly positively correlated with local VOC frequency for the Alpha SARS-CoV-2 variant and had a similar magnitude to that expected on the basis of its increased transmissibility, as estimated by other recent studies. All of these provide confidence in the model's nowcasting ability.

While the method is highly valuable and performs well on the UK data, the model still has a limitation: despite the different validation procedures, a true external validation on data from other countries could not be performed, as randomized testing is surprisingly and unfortunately rare in other countries. The UK was the

first to roll out regular randomized testing and has shown its great value for epidemic nowcasting and for public health responses at large. It will be interesting to see the use of this framework in other countries, if and when they roll out their own randomized testing campaigns. This work highlights the potential of randomized testing as an important tool in epidemic control and as a public health resource that has been largely undervalued and underused thus far in the pandemic.

One of the most exciting aspects of the statistical framework developed by Nicholson et al. is that it allows modular interoperability, for example it can be combined with various other epidemiological models and can be adjusted for different inputs (tests). Testing has seen rapid technological advancement over the months of the pandemic, and a plethora of different testing schemes have been added to the public health toolbox. It would be valuable to combine all of these into a unified causal framework to achieve data fusion⁵ at scale. A few testing tools that could be combined include population-wide wastewater testing and other pooled testing schemes; different types of rapid tests with varying test characteristics; semi-randomized surveillance testing such as that done regularly in hospitals, in care homes and at airports; and various digital epidemiological surveillance technologies⁶. Incorporating one-time full-population testing campaigns such as those done in Liverpool or Slovakia⁷ would also be valuable. Surely, all of these inputs contain

useful information, and the challenge is determining how to properly fuse them together. Such a framework will be useful not only for SARS-CoV-2 nowcasting but also for epidemic surveillance of other pathogens now and in the future. Therefore, the present framework provides a good proof of concept for a complete integrative model that has the ability to efficiently assimilate different data sources from different technologies and with different biases. This has the potential to greatly assist the world-wide public health response to both the current pandemic and future ones. □

Hagai Rossman^{1,2} and Eran Segal ^{1,2} ✉

¹Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. ²Department of Molecular Cell Biology, Weizmann Institute of Science, Rehovot, Israel.

✉e-mail: eran.segal@weizmann.ac.il

Published online: 31 December 2021
<https://doi.org/10.1038/s41564-021-01035-2>

References

1. Accorsi, E. K. et al. *Eur. J. Epidemiol.* **36**, 179–196 (2021).
2. Nicholson, G. et al. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-021-01029-0> (2021).
3. Wu, J. T. et al. *Nat. Med.* **27**, 388–395 (2021).
4. Mercer, T. R. & Salit, M. *Nat. Rev. Genet.* **22**, 415–426 (2021).
5. Bareinboim, E. & Pearl, J. *Proc. Natl Acad. Sci. USA*. **113**, 7345–7352 (2016).
6. Budd, J. et al. *Nat. Med.* **26**, 1183–1192 (2020).
7. Pavelka, M. et al. *Science* **372**, 635–641 (2021).

Competing interests

The authors declare no competing interests.