

of evaporated ocean water into ice sheets or other freshwater reservoirs, which can affect sea level. For certain times during recent ice ages, sea-level records obtained from isotopic analyses of microfossils are inconsistent with records derived from corals⁸. Geibert *et al.* suggest that these inconsistencies could be explained by the proposed storage of large volumes of fresh water in the Arctic Ocean.

Various complications in the analysis will no doubt raise questions. Arctic sediments are notoriously hard to date owing to the lack of microfossils, and because sedimentation rates varied^{2,3}. It is therefore uncertain whether the ²³⁰Th_{ex}-deficient intervals in the cores were produced at exactly the same times at all sites across the ocean basins. Moreover, the authors had to correct their data to account for ²³⁰Th that was produced from uranium decay in sediment grains, rather than in the water column⁵, and this contributed to the uncertainty in measured ²³⁰Th_{ex}. These corrections become proportionally more important for older sediments because ²³⁰Th_{ex} itself decays away; thorium decay also limits the time span over which the method can be used to investigate Arctic salinity. Finally, no freshwater fauna have been identified in the sediments concerned, so direct evidence of freshwater intrusion into deep Arctic basins remains to be found.

However, the various absences – of ²³⁰Th_{ex}, of microfossils and biological productivity, and of elements such as sulfur, which partly derive from salinity in marine sediments⁹ – suggest exciting avenues for future research. Computational modelling of Arctic Ocean circulation and ice-sheet behaviour will be needed to determine realistic estimates of the circulation and freshwater run-off from land that could produce a basin filled with fresh water. Further geochemical and fossil analyses might help to support or challenge the assertion that the Arctic Ocean could have been fresh. Geibert and colleagues' innovative use of ²³⁰Th might spur a re-evaluation of what is possible in the Arctic Ocean, and of how dramatically this region can change.

Sharon Hoffmann is a palaeoceanographer based in Wilmington, North Carolina, USA. e-mail: sshoffma@gmail.com

1. Arctic Climate Impact Assessment. *Impacts of a Warming Arctic* (Cambridge Univ. Press, 2004).
2. Jakobsson, M. *et al.* *Quat. Sci. Rev.* **92**, 40–67 (2014).
3. Polyak, L. *et al.* *Glob. Planet. Change* **68**, 5–17 (2009).
4. Geibert, W., Matthiessen, J., Stimac, I., Wollenburg, J. & Stein, R. *Nature* **590**, 97–102 (2021).
5. Henderson, G. M. & Anderson, R. F. *Rev. Mineral. Geochem.* **52**, 493–531 (2003).
6. Costa, K. M. *et al.* *Paleoceanogr. Paleoclimatol.* **35**, e2019PA003820 (2020).
7. Jakobsson, M. *et al.* *Nature Commun.* **7**, 10365 (2016).
8. Rohling, E. J. *et al.* *Quat. Sci. Rev.* **176**, 1–28 (2017).
9. Jørgensen, B. B. & Kasten, S. in *Marine Geochemistry* (eds Schulz, H. D. & Zabel, M.) 271–309 (Springer, 2006).

Coronavirus

An estimation of undetected COVID cases

Jeffrey Shaman

Most SARS-CoV-2 infections are undocumented. French health-care records and modelling were used to assess the rate of documentation of COVID-19 cases. The findings highlight the need for improved identification of infections. **See p.134**

The ongoing COVID-19 pandemic has stretched health-care systems, disrupted economies and frayed the fabric of societies around the globe. Nation states have responded differently to the crisis, and only a few countries have managed to achieve continued control of coronavirus transmission. On page 134, Pullano *et al.*¹ report their study of SARS-CoV-2 infections in France during late spring and early summer of 2020. Their analysis came after the peak of the first wave of COVID-19 cases there, but before the increase again in the autumn, which led to a higher second wave of cases. The results offer some important lessons.

Three key characteristics of the SARS-CoV-2 coronavirus have made it particularly challenging and disruptive. First, it is a newly emerged virus to which most people have little or no pre-existing immunity. Second, most infections are not documented by health-care systems; many individuals with these undocumented infections have no or only mild symptoms and can unknowingly spread the virus². Moreover, the individuals with infections that are eventually documented typically become contagious before symptoms appear^{2,3}. As a result, most of the viral transmission within and between communities is driven by a combination of pre-symptomatic individuals who will later develop symptoms, those who never develop symptoms (asymptomatic individuals) and those with mild symptoms who do not seek medical attention. Third, despite the preponderance of undocumented infections, the total number of global infections (documented and undocumented) is so high – with hundreds of millions infected so far – that the relatively small fraction of infections that result in severe outcomes or death⁴ already numbers in the millions.

To understand and control the pandemic, it is crucial to estimate how widespread the disease is within communities. One means of determining infection prevalence is provided by a measurement called the ascertainment rate – the fraction of all infections that are documented in the health-care system as

confirmed cases. A higher ascertainment rate signifies a greater capacity to identify infections and thus possibly control the onward spread of the virus through the isolation of infected individuals, along with other interventions. Pullano and colleagues present two independently generated estimates of new COVID-19 symptomatic cases in France during a seven-week period from mid-May to June (generating both these estimates using data and inference approaches), to quantify the ascertainment rate for symptomatic infections. They find a substantial discrepancy between their estimate of the number of detected symptomatic cases and of the total number of detected and undetected symptomatic cases (Fig. 1), revealing a low ascertainment rate. The findings suggest that the overall testing and control system in place was inadequate to contain the virus successfully in this country of around 65 million people.

Pullano and colleagues derived their first estimate of confirmed symptomatic cases of COVID-19 from a nationwide database, which was designed to track and record SARS-CoV-2 test results. This database was launched in May 2020, at the end of the spring lockdown in France. The information it contains includes routine virological testing and symptom reports for: health-care personnel; elderly individuals in nursing homes; residents of long-term care facilities; and patients hospitalized for any reason, as well as data on the tracing and testing of contacts of confirmed COVID-19 cases. In addition to these surveillance measures, any individual who showed COVID-19 symptoms could obtain diagnostic testing if referred by their doctor. By accounting for asymptomatic and pre-symptomatic cases, as well as for delays from symptom onset to testing, Pullano *et al.* obtained an estimate of the number of people with confirmed COVID-19 in whom symptoms developed during the seven-week study period.

For their second estimate, Pullano and colleagues used a mathematical model, hospitalization records and a maximum-likelihood

estimation method to infer all incidences of SARS-CoV-2 (both documented and undocumented) in which symptom onset in infected individuals occurred during the study period. Their model-inference approach enables the estimation of a broad suite of epidemiological conditions, including the total number of symptomatic infections. Such estimates have uncertainties depending on the quality of the data available and whether the assumptions and form of the model truly represent the dynamics of SARS-CoV-2 transmission. By systematically testing the sensitivity of their findings to the assumptions they made, Pullano *et al.* provide extra confidence in their inferences and constrain the boundaries of the uncertainty for their estimated value for the total number of symptomatic infections.

Using the surveillance database, Pullano *et al.* estimate that there were more than 14,000 confirmed symptomatic cases with symptom onset during the study period. However, using the model-inference approach, the authors instead suggest a much higher estimate – of just under 104,000 symptomatic infections in total (both confirmed and those never documented). This finding paints a worrying overall picture. The estimate from the model-inference approach suggests that 86% of symptomatic infections were undetected, despite the low infection levels that followed the spring lockdown and a centralized, coordinated testing and tracing programme. Pullano *et al.* rightly note that although the detection rate rose during the course of the seven-week period studied, the vast majority of symptomatic infections were nevertheless undetected.

Pullano and colleagues did not include asymptomatic infections when making their calculation of the ascertainment rate. If the authors' estimated asymptomatic infection rate is included, making the corresponding calculation suggests that only approximately 1 in 12 SARS-CoV-2 infections was identified during the study period. Other studies^{2,3} have similarly estimated a substantial under-counting of total infections and symptomatic infections. Pullano *et al.* also use their findings to consider what the data reveal about shortcomings in COVID-19 surveillance efforts.

The testing and tracing systems implemented around the world vary, but they begin with 'index' cases, individuals without known infectious contacts who are tested because of their work or living circumstances, the presentation of symptoms, or in some instances simply because of a desire to be tested. By requiring that most of the population tested obtain a referral from their physician, France might have disadvantaged communities that have more-limited access to health care, and reduced testing rates. Pullano and colleagues also report that surveillance data suggest that less than one-third of individuals in France

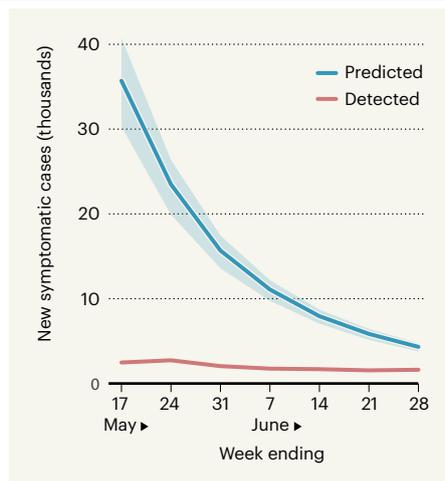


Figure 1 | COVID-19 cases in France. Pullano *et al.*¹ present their analysis of the number of new cases of COVID-19 in which individuals developed symptoms during a seven-week period, from 11 May to 28 June 2020. This corresponded to a time when infections were falling after the end of a national lockdown. Using a national database, the authors estimated the number of confirmed symptomatic cases detected in the health-care system at the end of each week (red line). They used mathematical modelling and assessment of hospital records to predict the total number of new COVID-19 cases, including those not recorded in the medical system (blue line, shading indicates the 95% confidence interval). The results suggest that a high proportion of cases were undetected. (Graph based on Fig. 3a of ref. 1.)

who had symptoms of COVID-19 consulted a physician.

In such circumstances, the prospects for continued viral control were probably compromised. France subsequently experienced a particularly devastating wave of COVID-19 cases beginning in August and peaking in early November. The policies and behaviours that limited virus spread during early summer were insufficient to control the virus during late summer and autumn, when some people went on holiday or grew tired of social distancing, schools and universities opened, and the weather turned colder. This last factor resulted in greater virus survival^{6,7} and people spending more time indoors and in closer proximity to others, compared with during the summer⁸.

Some countries, such as South Korea, Thailand and Vietnam, have managed to use testing and tracing in conjunction with other control strategies to keep ascertainment rates high and the virus largely in check (see go.nature.com/3qkm2bz). These successes indicate that testing and tracing programmes can be effective, provided that the prevalence of infections does not exceed the capacity to test for and document infections, to identify and trace contacts extensively, and to isolate and quarantine contagious and potentially contagious individuals. However, testing and tracing alone

are probably insufficient to achieve sustained COVID-19 control⁹. Instead, political and public effort is also needed to ensure continued compliance with interventions to limit infections, including the wearing of face masks, social distancing and the restriction of large gatherings¹⁰. Many countries, as a result of leadership failures, cultural or institutional barriers, or simple fatigue, have failed in their efforts to achieve or maintain control of the virus.

As COVID-19 continues to spread in France and elsewhere, it is imperative that ascertainment rates be monitored, either using model-inference approaches² or by widespread testing of individuals for antibodies to SARS-CoV-2, which provides an estimate of the proportion of a population previously infected with the virus¹¹. Pullano and colleagues' findings support the idea that efforts should be made to track rates of under-detection and to better quantify the overall number of infections. Such information is crucial for accurately evaluating control measures, so that the need for stepped-up or alternative interventions can be identified, and for developing vaccine-deployment strategies. (For example, understanding the infection profile of a region provides useful information if planning a cluster-busting vaccination approach.) This latter issue will be particularly crucial given the timescales of vaccine production and distribution, and the time needed after vaccination for a protective immune response to develop.

Pullano *et al.* use the ascertainment rate to quantify under-detection in France, and their results highlight how such under-detection, if recognized, could be used as a motivation for implementing better pandemic surveillance and control measures. The authors also show the value of mathematical modelling for assessing both population-scale disease prevalence and the effectiveness of disease-control policies.

Jeffrey Shaman is in the Department of Environmental Health Sciences, Columbia University Mailman School of Public Health, New York, New York 10032, USA.
e-mail: jlsh106@cumc.columbia.edu

1. Pullano, G. *et al.* *Nature* **590**, 134–139 (2021).
2. Li, R. *et al.* *Science* **368**, 489–493 (2020).
3. He, X. *et al.* *Nature Med.* **26**, 672–675 (2020).
4. Yang W. *et al.* *Lancet Infect. Dis.* [https://doi.org/10.1016/S1473-3099\(20\)30769-6](https://doi.org/10.1016/S1473-3099(20)30769-6) (2020).
5. Flaxman, S. *et al.* *Nature* **584**, 257–261 (2020).
6. Morris, D. H. *et al.* Preprint at [bioRxiv](https://doi.org/10.1101/2020.10.16.341883) <https://doi.org/10.1101/2020.10.16.341883> (2020).
7. Ma, Y., Pei, S., Shaman, J., Dubrow, R. & Chen, K. Preprint at [medRxiv](https://doi.org/10.1101/2020.11.13.20231472) <https://doi.org/10.1101/2020.11.13.20231472> (2020).
8. Qian, H. *et al.* *Indoor Air* <https://doi.org/10.1111/ina.12766> (2020).
9. Dighhe, A. *et al.* *BMC Med.* **18**, 321 (2020).
10. Hsiang, S. *et al.* *Nature* **584**, 262–267 (2020).
11. Pollán, M. *et al.* *Lancet* **396**, 535–544 (2020).

The author declares competing financial interests: see go.nature.com/3qlcvro for details.

This article was published online on 21 December 2020.