QCOVID in Scotland: time to recalibrate our expectations?

Rishi K Gupta [iD],[1] Maarten van Smeden[2]

The field of prediction modelling has exploded during the COVID-19 pandemic, with countless studies seeking to derive and/or validate models to assist diagnosis of SARS-CoV-2 or to predict clinical outcomes of infection. While the vast majority of reports have been of a poor standard,[1] there have been some examples of higher quality work, including the QCOVID prognostic model.[2] Whereas most prediction models are intended to predict a single outcome, QCOVID seeks to predict the risk of acquiring and then either dying from or being hospitalised with SARS-CoV-2 in the general population. The model was originally derived using data from 6.1 million people in England during the first pandemic wave (January–April 2020) and showed promising performance in a recent analysis by the Office for National Statistics over a similar time period.[3] QCOVID is implemented online,[4] although stipulated for research rather than clinical use, and provides distinct probabilities of COVID-19 hospitalisation and COVID-19 mortality.

Simpson et al report an external validation of QCOVID using data from Scotland over two periods (1 March 2020 to 30 April 2020 and 1 May 2020 to 30 June 2020), designed to test the performance of the algorithms during time intervals that aligned with the original QCOVID derivation and temporal validation cohorts.[5] The authors use a population-based approach, including data from primary care facilities covering an impressive 99% of the Scottish population. The data provided by Simpson et al can help us to address a series of important questions regarding the performance of QCOVID in Scotland.

First, are the predictions from QCOVID accurate? To address this, the authors evaluated the model's 'calibration', which assesses the agreement between the model's predicted risk estimates and the observed risk in the population. To do this, they defined 20 groups of equal size based on predicted risk and compared the average predicted versus observed risk in each group. Calibration was judged to be reasonable overall during the first study period. However, there was marked overestimation of risk (where predicted risks were systematically higher than actual population risks) during the second period among males and females, for both outcomes of COVID-19 hospitalisation and COVID-19 mortality. This miscalibration was reflected in the 'observed:-expected ratios', calculated by dividing the overall number of observed outcome events by the number predicted by QCOVID. These ratios ranged from 0.26 to 1.07 during the second period, indicating miscalibration by a factor of up to 4 overall. This miscalibration likely reflects that QCOVID seeks to predict acquisition of SARS-CoV-2 followed by adverse outcomes, as single absolute probabilities. Such an approach inherently assumes a constant risk of infection over time and across regions—an assumption that we know is not true from epidemiological data. In reality, national lockdown was introduced in Scotland on 24 March 2020, leading to a marked reduction in SARS-CoV-2 transmission by mid-late April 2020. It is therefore unsurprising that the QCOVID algorithm, trained using data from January–April 2020, overestimated risks for May–June 2020.

So how can one deal with the challenging issue of constantly changing risks over time? One approach, termed 'temporal recalibration', seeks to tailor risk estimates for any given time period by adjusting for the baseline risk (or the overall number of outcome events) during that interval.[6] The authors conducted temporal recalibration for the second time period and, unsurprisingly, demonstrate improved calibration. However, this approach requires prior knowledge of the number of outcome events during the time period of interest. But how can we do this when we do not know the future baseline risk? Recalibration to the most recent time period available could be one approach, although validation in future intervals is still required. Recalibration using contemporary local proxy measures of transmission could also be considered to enable geographically tailored estimates. Nevertheless, Simpson et al's findings of miscalibration clearly demonstrate a need for ongoing recalibration of QCOVID, although full model refitting (where all coefficients are re-estimated and new predictors such as vaccination status may be added) could also be considered. Even still, the absolute risk estimates provided by QCOVID should be interpreted with caution, not least because of ever changing risks of infection over time and across geographical regions.

So if the absolute risks provided by QCOVID might not be accurate, then should we still use the tool in practice? If the intention of QCOVID is to inform the order in which interventions are applied, for example, prioritising vaccine rollout, then perhaps the ranking of predictions across the population matters more than the absolute values. Simpson et al assessed the accuracy of this ranking for QCOVID by quantifying the model's 'discrimination'—the degree to which predictions differentiated people who did and did not meet the outcomes of interest. One measure of discrimination for a time to event model such as QCOVID is Harrell's C statistic, which can range from 0 to 1, with 1 being perfect. The authors report Harrell's C statistics of 0.93–0.96 for COVID-19 mortality and 0.79–0.83 for COVID-19 hospitalisation. However, we know that age is a very strong predictor of poor outcomes in COVID-19. Notably, analyses early in the pandemic showed that multivariable models added no clinical utility over and above age as a single predictor to predict death among adults hospitalised with COVID-19.[7] Thus, it is challenging to contextualise the Harrell's C statistics provided for QCOVID in the absence of equivalent performance measures when using simpler approaches such as age alone, particularly given that COVID-19 mortality and COVID-19 hospitalisation are infrequent events in the general population, with markedly skewed risk distributions. It therefore remains unclear how much incremental discriminatory value QCOVID adds to a simplified approach that uses age (perhaps accompanied by a specified list of severe comorbidities), to prioritise interventions such as vaccine rollout.

External validation studies, where a model is tested in a population outside of its original derivation, are an integral component of the evaluation of prediction models to establish their

[1]Institute for Global Health, University College London, London, UK
[2]Julius Center for Health Sciences and Primary Care, Utrecht University, Utrecht, The Netherlands

Correspondence to Dr Rishi K Gupta, University College London Institute for Global Health, London, WC1E 6BT, UK; r.gupta@ucl.ac.uk

potential generalisability in new populations. Overall, the authors report that QCOVID performed similarly in this Scottish data set to the original English validation cohort. However, the analysis by Simpson and colleagues highlights a clear need for a robust and innovative plan to dynamically update the currently implemented version of QCOVID, particularly post-vaccine rollout. Additional evidence of QCOVID's clinical utility, over and above simpler approaches, to inform individual-level and wider policy decisions is also required.

Twitter Rishi K Gupta @Rishi_K_Gupta and Maarten van Smeden @MaartenvSmeden

Check for updates

Linked

► http://dx.doi.org/10.1136/thoraxjnl-2021-217580

ORCID iD
Rishi K Gupta http://orcid.org/0000-0002-6257-1285

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;325:m1328.
2 Clift AK, Coupland CAC, Keogh RH, *et al*. Living risk prediction algorithm (QCOVID) for risk of hospital admission and mortality from coronavirus 19 in adults: national derivation and validation cohort study. *BMJ* 2020;371:m3731.
3 Nafilyan V, Humberstone B, Mehta N, *et al*. An external validation of the QCovid risk prediction algorithm for risk of mortality from COVID-19 in adults: a national validation cohort study in England. *Lancet Digit Health* 2021;3:e425–33.
4 QCovidTM risk calculator. Available: https://qcovid.org/Calculation
5 Simpson CR, Robertson C, Kerr S, *et al*. External validation of the QCovid risk prediction algorithm for risk of COVID-19 hospitalisation and mortality in adults: national validation cohort study in Scotland. *Thorax* 2021:1–7.
6 Booth S, Riley RD, Ensor J, *et al*. Temporal recalibration for improving prognostic model development and risk predictions in settings where survival is improving over time. *Int J Epidemiol* 2020;49:1316–25.
7 Gupta RK, Marks M, Samuels THA, *et al*. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. *Eur Respir J* 2020;56:2003498.