

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



Vahé Nafilyan, Ben Humberstone, Nisha Mehta, Ian Diamond, Carol Coupland, Luke Lorenzi, Piotr Pawelek, Ryan Schofield, Jasper Morgan, Paul Brown, Ronan Lyons, Aziz Sheikh, Julia Hippisley-Cox



Summary

Background Public policy measures and clinical risk assessments relevant to COVID-19 need to be aided by risk prediction models that are rigorously developed and validated. We aimed to externally validate a risk prediction algorithm (QCovid) to estimate mortality outcomes from COVID-19 in adults in England.

Methods We did a population-based cohort study using the UK Office for National Statistics Public Health Linked Data Asset, a cohort of individuals aged 19–100 years, based on the 2011 census and linked to Hospital Episode Statistics, the General Practice Extraction Service data for pandemic planning and research, and radiotherapy and systemic chemotherapy records. The primary outcome was time to COVID-19 death, defined as confirmed or suspected COVID-19 death as per death certification. Two periods were used: (1) Jan 24 to April 30, 2020, and (2) May 1 to July 28, 2020. We assessed the performance of the QCovid algorithms using measures of discrimination and calibration. Using predicted 90-day risk of COVID-19 death, we calculated r^2 values, Brier scores, and measures of discrimination and calibration with corresponding 95% CIs over the two time periods.

Findings We included 34 897 648 adults aged 19–100 years resident in England. 26 985 (0·08%) COVID-19 deaths occurred during the first period and 13 177 (0·04%) during the second. The algorithms had good discrimination and calibration in both periods. In the first period, they explained 77·1% (95% CI 76·9–77·4) of the variation in time to death in men and 76·3% (76·0–76·6) in women. The D statistic was 3·761 (3·732–3·789) for men and 3·671 (3·640–3·702) for women and Harrell's C was 0·935 (0·933–0·937) for men and 0·945 (0·943–0·947) for women. Similar results were obtained for the second time period. In the top 5% of patients with the highest predicted risks of death, the sensitivity for identifying deaths in the first period was 65·94% for men and 71·67% for women.

Interpretation The QCovid population-based risk algorithm performed well, showing high levels of discrimination for COVID-19 deaths in men and women for both time periods. QCovid has the potential to be dynamically updated as the pandemic evolves and, therefore, has potential use in guiding national policy.

Funding UK National Institute for Health Research.

Copyright © 2021 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY-NC-ND 4.0 license.

Introduction

The first cases of SARS-CoV-2 infection were reported in the UK on Jan 24, 2020, with the first COVID-19 death on Feb 28, 2020. As of May 11, 2021, over 127 000 deaths from COVID-19 have occurred in the UK, and over 3 million deaths globally. Emerging evidence throughout the course of the COVID-19 pandemic, initially from case series and then from cohorts of individuals with confirmed SARS-CoV-2 infection, has shown associations of age, sex, certain comorbidities, ethnicity, and obesity with adverse COVID-19 outcomes such as hospitalisation and death.^{1–8} A growing knowledge base now exists regarding risk factors for severe COVID-19. As many countries are re-introducing lockdown measures and vaccination programmes have started being rolled out, the opportunity exists to develop more nuanced guidance⁹ that is based on predictive algorithms to inform

risk-management decisions. Improved knowledge of individuals' risks could also help guide decisions on managing occupational risk and in the targeting of vaccines to those most at risk. Although several risk-prediction models have been developed, a systematic review¹⁰ found that most models have high risk of bias and that their reported performance is optimistic.

The use of primary care datasets such as QResearch, with linkage to registries such as death records and hospital admissions data, represents an innovative approach to clinical risk prediction modelling for COVID-19, which has successfully been developed, validated, and implemented in the UK National Health Service (NHS) over the past 10 years.^{11–13} The method provides accurately coded, individual-level data for many people representative of the national population. This approach was used to develop the QCovid prediction

Lancet Digit Health 2021;
3: e425–33

Published Online
May 25, 2021
[https://doi.org/10.1016/S2589-7500\(21\)00080-7](https://doi.org/10.1016/S2589-7500(21)00080-7)

Office for National Statistics, Newport, UK (V Nafilyan PhD, B Humberstone MRes, Prof I Diamond PhD, L Lorenzi MSc, P Pawelek MSc, R Schofield BSc, J Morgan BSc, P Brown); Office of the Chief Medical Officer, Department of Health & Social Care, London, UK (N Mehta MD); Division of Primary Care, School of Medicine, University of Nottingham, Nottingham, UK (Prof C Coupland PhD); National Centre for Population Health and Wellbeing Research, Swansea University Medical School, Swansea, UK (Prof R Lyons MD); Usher Institute, University of Edinburgh, Edinburgh, UK (Prof A Sheikh MD); Nuffield Department of Primary Health Care Sciences, University of Oxford, Oxford, UK (Prof J Hippisley-Cox MD)

Correspondence to:
Dr Vahé Nafilyan, Office for National Statistics, Newport, UK
vahe.nafilyan@ons.gov.uk

For COVID-19 statistics see
<https://coronavirus.jhu.edu/map.html>

Research in context

Evidence before this study

We searched PubMed for articles about the validation of existing predictive models, using the search terms “COVID-19”, “risk”, “prediction”, and “validation”, focusing on studies published between March 1 and Dec 31, 2020. No study had validated the QCovid risk prediction algorithm. Public policy measures and clinical risk assessments relevant to COVID-19 need to be aided by rigorously developed and validated risk prediction models. A recent living systematic review of published risk prediction models for COVID-19 found that most models were subject to a high risk of bias with optimistic reported performance, raising concern that these models might be unreliable when applied in practice. A population-based risk prediction model, QCovid risk prediction algorithm, has been developed to identify adults at high risk of serious COVID-19 outcomes, which overcomes many of the limitations of previous tools.

Added value of this study

Commissioned by the Chief Medical Officer for England, we validated the novel clinical risk prediction model QCovid to identify risks of short-term severe outcomes due to COVID-19. We used national linked datasets from general practice, death registry data, and Hospital Episode Statistics data for a population-representative sample of more than 34 million adults. The risk models had excellent discrimination in men and women and were well calibrated. QCovid represents a new, evidence-based opportunity for population risk stratification.

Implications of all the available evidence

QCovid has the potential to support public health policy, from enabling shared decision making between clinicians and patients in relation to health and work risks, to targeted recruitment for clinical trials and prioritisation of vaccination, for example.

models,¹⁴ drawing on the rich phenotyping of individuals with demographic, medical, and pharmacological predictors to allow robust statistical modelling and assessment. Such linked datasets have a track record for the development and assessment of established clinical risk models including for cardiovascular disease,¹¹ diabetes (either type 1 or type 2),¹³ and mortality.¹² Although QCovid predicts both COVID-19-related hospital admission and death, the aim of this analysis was to validate the outcome that estimates the risks of becoming infected and subsequent death due to COVID-19 in a large national cohort.

Methods

Study design and data sources

The Chief Medical Officer for England asked the New and Emerging Respiratory Virus Threats Advisory Group to develop and validate a clinical risk prediction model for COVID-19 in line with the emerging evidence. The resulting QCovid model was developed and validated using the QResearch database and reported in accordance with Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis and The REporting of studies Conducted using Observational Routinely-collected health Data guidelines and with input from a patient advisory group. This paper reports the validation of the model on an independent data source.

We undertook a validation cohort study of individuals aged 19–100 years using the UK Office for National Statistics (ONS) Public Health Linked Data Asset. This dataset is based on the 2011 census in England, linked at an individual level using the NHS number to mortality records, Hospital Episode Statistics, and the General Practice Extraction Service (GPES) data for pandemic planning and research. To obtain NHS numbers, the

2011 census was linked to the 2011–13 NHS patient registers using deterministic and probabilistic matching, with an overall linkage rate of 94.6%. We excluded patients (approximately 13.1%) who did not have a valid NHS number or were not found in primary care records. To validate the QCovid algorithm, we further linked radiotherapy and systemic chemotherapy records on the basis of NHS number. The ONS Public Health Linked Data Asset includes data on most patients used to develop the QCovid algorithm but also includes patients registered with practices using information technology systems other than Egton Medical Information Systems (also known as EMIS), such as The Phoenix Partnership (also known as TPP; used by 35% of general practitioner [GP] practices in England).

We identified a cohort of all individuals aged 19–100 years who were enumerated at the 2011 census and registered alive and resident in England on Jan 24, 2020. Patients entered the cohort on Jan 24, 2020 (date of first confirmed COVID-19 case in UK) and were followed up until they had the outcome of interest or July 28, 2020, which is the date up to which linked data were available at the time of the analysis. This date also extends the period of observation beyond the original QCovid study. We divided the study period into two time periods: Jan 24, to April 30, 2020, and May 1, to July 28, 2020.

Outcomes

The primary outcome was death involving COVID-19 (either in hospital or out of hospital), defined as confirmed or suspected COVID-19 death as identified by two codes of the tenth revision of the International Classification of Diseases (U07.1 or U07.2) recorded on the death certification. The time-at-risk was calculated from the beginning of each period (Jan 24, 2020, or May 1, 2020).

Panel: Predictor variables used to validate the QCovid model

- Age in years (continuous)
- Townsend deprivation score (continuous)
- Accommodation (neither homeless nor care home vs care home or nursing home)
- Ethnicity in ten categories (Bangladeshi, Black African, Black Caribbean, Chinese, Indian, Mixed, Pakistani, White British, White other, Other)
- Body-mass index (kg/m²)
- Chronic kidney disease* (no chronic kidney disease, stage 3, stage 4, or stage 5)
- Learning disability (no learning disability, Down Syndrome, or other learning disability)
- Chemotherapy in the past 12 months (chemotherapy group A, B, or C, based on the risk of grade 3 or 4 febrile neutropenia [Common Terminology Criteria for Adverse Events version 4] or lymphopenia)
- Respiratory cancer
- Radiotherapy in the past 6 months
- Solid organ transplant
- Prescribed immunosuppressant medication by general practitioner
- Prescribed leukotriene or long-acting β_2 agonists
- Prescribed regular prednisolone
- Diabetes[†]
- Chronic obstructive pulmonary disease
- Asthma
- Rare pulmonary diseases
- Pulmonary hypertension or pulmonary fibrosis
- Coronary heart disease
- Stroke
- Atrial fibrillation
- Congestive cardiac failure
- Venous thromboembolism
- Peripheral vascular disease
- Congenital heart disease
- Dementia
- Parkinson's disease
- Epilepsy
- Rare neurological conditions
- Cerebral palsy
- Severe mental illness (bipolar disorder, schizophrenia, or severe depression)
- Osteoporotic fracture
- Rheumatoid arthritis or systemic lupus erythematosus
- Cirrhosis of the liver

*Patients with stage 5 chronic kidney disease were assigned the coefficient for stage 5 disease without transplant nor dialysis. [†]For the validation of the QCovid risk model, all patients with diabetes were assigned the coefficient type 2 diabetes.

Predictor variables

We derived pre-existing conditions and demographic characteristics using the same definitions as those used to develop the QCovid algorithm. Demographic characteristics were taken from the 2011 census.

	Overall	Period 1 (Jan 24, to April 30, 2020)	Period 2 (May 1, to July 31, 2020)
Overall	34 897 648	26 985	13 177
Sex			
Female	18 297 773 (52.43%)	11 651 (43.18%)	6560 (49.78%)
Male	16 599 875 (47.57%)	15 334 (56.82%)	6617 (50.22%)
Age, years	51.09 (18.76)	79.98 (11.63)	82.13 (10.79)
Age group, years			
19–29	5 601 475 (16.05%)	44 (0.16%)	13 (0.10%)
30–39	5 268 030 (15.10%)	116 (0.43%)	30 (0.23%)
40–49	5 625 225 (16.12%)	364 (1.35%)	125 (0.95%)
50–59	6 435 204 (18.44%)	1196 (4.43%)	400 (3.04%)
60–69	5 185 917 (14.86%)	2727 (10.11%)	962 (7.30%)
70–79	4 225 729 (12.11%)	6280 (23.27%)	2695 (20.45%)
80–89	2 093 545 (6.00%)	10 841 (40.17%)	5580 (42.35%)
≥90	462 523 (1.33%)	5417 (20.07%)	3372 (25.59%)
Geographical region			
East Midlands	3 137 521 (8.99%)	1979 (7.33%)	1372 (10.41%)
East of England	3 987 067 (11.43%)	2549 (9.45%)	1456 (11.05%)
London	4 662 731 (13.36%)	5403 (20.02%)	956 (7.26%)
North east	1 755 316 (5.03%)	1429 (5.30%)	931 (7.07%)
North west	4 643 947 (13.31%)	4289 (15.89%)	2411 (18.30%)
South east	5 818 470 (16.67%)	4005 (14.84%)	2118 (16.07%)
South west	3 674 549 (10.53%)	1657 (6.14%)	745 (5.65%)
West Midlands	3 643 447 (10.44%)	3284 (12.17%)	1497 (11.36%)
Yorkshire and the Humber	3 574 600 (10.24%)	2390 (8.86%)	1691 (12.83%)
Ethnicity			
Bangladeshi	258 053 (0.74%)	179 (0.66%)	29 (0.22%)
Black African	520 547 (1.49%)	398 (1.47%)	62 (0.47%)
Black Caribbean	374 982 (1.07%)	732 (2.71%)	124 (0.94%)
Chinese	185 966 (0.53%)	107 (0.40%)	27 (0.20%)
Indian	931 247 (2.67%)	800 (2.96%)	216 (1.64%)
Mixed	551 567 (1.58%)	184 (0.68%)	67 (0.51%)
Other	835 506 (2.39%)	590 (2.19%)	130 (0.99%)
Pakistani	679 062 (1.95%)	426 (1.58%)	123 (0.93%)
White British	28 845 085 (82.66%)	22 462 (83.24%)	12 018 (91.20%)
White other	1 715 633 (4.92%)	1107 (4.10%)	381 (2.89%)
Townsend deprivation quintile			
1 (most affluent)	7 491 652 (21.47%)	4993 (18.50%)	2842 (21.57%)
2	7 738 292 (22.17%)	5326 (19.74%)	2967 (22.52%)
3	6 834 804 (19.58%)	5111 (18.94%)	2647 (20.09%)
4	6 467 204 (18.53%)	5365 (19.88%)	2472 (18.76%)
5 (most deprived)	6 366 096 (18.24%)	6190 (22.94%)	2249 (17.07%)
Accommodation			
Neither homeless nor care home	34 667 007 (99.34%)	19 995 (74.10%)	9039 (68.60%)
Care home or nursing home	230 641 (0.66%)	6990 (25.90%)	4138 (31.40%)
Body-mass index, kg/m ²			
<18.5	393 928 (1.13%)	983 (3.64%)	614 (4.66%)
18.5 to <25	6 658 276 (19.08%)	5776 (21.40%)	2965 (22.50%)
25 to <30	6 661 721 (19.09%)	5552 (20.57%)	2385 (18.10%)
≥30	5 661 007 (16.22%)	5540 (20.53%)	2066 (15.68%)
Not recorded	15 522 716 (44.48%)	9134 (33.85%)	5147 (39.06%)

(Table 1 continues on next page)

	Overall	Period 1 (Jan 24, to April 30, 2020)	Period 2 (May 1, to July 31, 2020)
(Continued from previous page)			
Chronic kidney disease			
No chronic kidney disease	34 392 544 (98.55%)	24 425 (90.51%)	11 939 (90.60%)
Stage 3	436 595 (1.25%)	1820 (6.74%)	914 (6.94%)
Stage 4	45 638 (0.13%)	452 (1.68%)	205 (1.56%)
Stage 5	22 871 (0.07%)	288 (1.07%)	119 (0.90%)
Learning disability			
No learning disability	34 393 288 (98.55%)	25 300 (93.76%)	12 386 (94.00%)
Learning disability	490 357 (1.41%)	1616 (5.99%)	*
Down Syndrome	14 003 (0.04%)	69 (0.26%)	*
Chemotherapy†			
No chemotherapy in past 12 months	34 776 317 (99.65%)	26 472 (98.10%)	12 908 (97.96%)
Chemotherapy group A	38 956 (0.11%)	128 (0.47%)	62 (0.47%)
Chemotherapy group B	76 763 (0.22%)	339 (1.26%)	180 (1.37%)
Chemotherapy group C	5612 (0.02%)	46 (0.17%)	27 (0.20%)
Cancer and immunosuppression			
Blood cancer	336 990 (0.97%)	897 (3.32%)	465 (3.53%)
Respiratory cancer	9720 (0.03%)	142 (0.53%)	66 (0.50%)
Radiotherapy in past 6 months	56 252 (0.16%)	174 (0.64%)	100 (0.76%)
Solid organ transplant	3488 (0.01%)	26 (0.10%)	*
Prescribed immunosuppressant medication by GP	7237 (0.02%)	20 (0.07%)	*
Prescribed leukotriene or LABA	2 362 855 (6.77%)	4956 (18.37%)	2319 (17.60%)
Prescribed regular prednisolone	404 467 (1.16%)	2124 (7.87%)	1028 (7.80%)
Other comorbidities			
Diabetes‡	3 087 792 (8.85%)	8700 (32.24%)	3650 (27.70%)
COPD	1 053 783 (3.02%)	3814 (14.13%)	1809 (13.73%)
Asthma	4 382 954 (12.56%)	3344 (12.39%)	1504 (11.41%)
Rare pulmonary diseases	373 807 (1.07%)	1707 (6.33%)	734 (5.57%)
Pulmonary hypertension or pulmonary fibrosis	127 760 (0.37%)	1158 (4.29%)	502 (3.81%)
Coronary heart disease	1 549 243 (4.44%)	5946 (22.03%)	2861 (21.71%)
Stroke	902 277 (2.59%)	5086 (18.85%)	2685 (20.38%)
Atrial fibrillation	1 096 209 (3.14%)	5237 (19.41%)	2894 (21.96%)
Congestive cardiac failure	545 617 (1.56%)	3739 (13.86%)	1830 (13.89%)
Venous thromboembolism	8878 (0.03%)	35 (0.13%)	*
Peripheral vascular disease	303 118 (0.87%)	1588 (5.88%)	771 (5.85%)
Congenital heart disease	359 (<0.01%)	*	0
Dementia	414 540 (1.19%)	8293 (30.73%)	4699 (35.66%)
Parkinson's disease	113 647 (0.33%)	1021 (3.78%)	573 (4.35%)
Epilepsy	405 047 (1.16%)	797 (2.95%)	387 (2.94%)
Rare neurological conditions	27 583 (0.08%)	149 (0.55%)	48 (0.36%)
Cerebral palsy	4350 (0.01%)	31 (0.11%)	*
Severe mental illness	6 574 526 (18.84%)	5341 (19.79%)	2541 (19.28%)
Osteoporotic fracture	29 153 (0.08%)	194 (0.72%)	92 (0.70%)
Rheumatoid arthritis or SLE	315 431 (0.90%)	696 (2.58%)	369 (2.80%)
Cirrhosis of the liver	81 753 (0.23%)	241 (0.89%)	114 (0.87%)

Data are n (%) or mean (SD). COPD=chronic obstructive pulmonary disease. GP=general practitioner. LABA=long-acting β2 agonist. SLE=systemic lupus erythematosus. *Represents values that have been suppressed due to small numbers (ie, <5). †Groups based on the risk of grade 3 or 4 febrile neutropenia (Common Terminology Criteria for Adverse Events version 4) or lymphopenia. ‡Included patients with either type 1 or type 2 diabetes.

Table 1: Demographic and medical characteristics for the validation cohort and those who died with COVID-19 in the two time periods

For comorbidities, we used data from Jan 1, 2015, to Dec 31, 2019. For body-mass index (BMI), we used the latest recorded value up to December 31, 2019.

The primary care records used in the ONS Public Health Linked Data Asset were based on an existing GPES dataset, which included many but not all of the relevant clinical codes used to develop the QCovid algorithm. Nonetheless, we derived data on most of the pre-existing conditions. However, we could not identify patients who had a solid organ or bone marrow transplant in the past 6 months, those on kidney dialysis or who had received a kidney transplant, or those with sickle cell disease or severe combined immunodeficiency syndrome. Similarly, we could not distinguish between patients with type 1 or type 2 diabetes. Variables used to validate the QCovid algorithm are listed in the panel.

Model validation

We fitted an imputation model to replace missing values for BMI, using predicted values from linear regression models stratified by sex. Predictors included all predictor variables in the QCovid algorithm, interacted with age.

We applied the QCovid risk equations (version 1), which are reported in the study that developed the algorithm,¹⁴ to men and women in the validation dataset. For conditions that we could not identify, we could not apply the coefficients from the QCovid risk equations. All patients with diabetes were assigned the coefficient for type 2 diabetes. Patients with stage 5 chronic kidney disease were assigned the coefficient for stage 5 chronic kidney disease without transplant nor dialysis.

Using predicted 90-day risk of COVID-19 death, we calculated *r*² values,²¹ Brier scores, and measures of discrimination and calibration^{22,23} with corresponding 95% CIs over the two time periods. *r*² values refer to the proportion of variation in survival time explained by the model. Brier scores measure predictive accuracy where lower values indicate better accuracy.²⁴ The D statistic is a discrimination measure that quantifies the separation in survival between patients with different levels of predicted risks and Harrell's C statistic is a discrimination metric that quantifies the extent to which people with higher risk scores have earlier events. Model calibration was assessed in the two time periods by comparing mean predicted risks with observed risks by 20ths of predicted risk. Observed risks were derived in each of the 20 groups using non-parametric estimates of the cumulative incidences.

The performance metrics were calculated in the whole cohort and in the following pre-specified subgroups: 5-year age-sex bands, ten ethnic groups, and within each quintile of the Townsend index, a measure of deprivation. We also estimated the performance metrics on a sample restricted to patients registered with practices using the TPP system and therefore not used at all to derive the algorithm. The code for this analysis is available on GitHub. We also derived the metrics for an alternative second period (May 1, 2020, to June 30, 2020), which was

	Period 1 (Jan 24, to April 30, 2020)		Period 2 (May 1, to July 31, 2020)	
	COVID-19 death in women	COVID-19 death in men	COVID-19 death in women	COVID-19 death in men
r^2 statistic	0.763 (0.760–0.766)	0.771 (0.769–0.774)	0.754 (0.750–0.757)	0.774 (0.769–0.777)
D statistic	3.671 (3.640–3.702)	3.761 (3.732–3.789)	3.579 (3.542–3.616)	3.782 (3.739–3.826)
Harrell's C statistic	0.945 (0.943–0.947)	0.935 (0.933–0.937)	0.956 (0.954–0.958)	0.944 (0.942–0.946)
Brier score	0.0018	0.0013	0.0007	0.0008

Data are estimate (95% CI).

Table 2: Performance of the risk models to predict risk of COVID-19 death in the validation cohort

the period used in the study that developed the algorithm. All analyses were done using R (version 3.5).

The ethics approval for the development and validation of QCovid was granted by the East Midlands–Derby Research Ethics Committee (18/EM/0400).

Role of the funding source

This study was funded by a grant from the UK National Institute for Health Research following a commission by the Chief Medical Officer for England whose office contributed to the development of the study question and facilitated access to relevant national datasets, contributed to interpretation of data, and drafting of the report.

Results

34 897 648 people in England aged 19–100 years met our inclusion criteria. Of the 40 136 597 people aged 19–100 years who were enumerated at the 2011 census and were alive on Jan 24, 2020, 2 071 521 (5.2%) people were excluded because they could not be linked to the 2011–13 NHS patient register and therefore did not have an NHS number. A further 3 167 428 (7.9%) people could not be linked to the GPES data, possibly because they migrated out of England and therefore were no longer registered with the NHS in England. Our data covered 80.0% of the population in England aged at least 19 years (appendix p 1). Coverage was lowest in London (4 662 731 [68.22%] of 6 834 636 people) and highest in Yorkshire and the Humber (3 574 600 [83.69%] of 4 271 381 people; appendix p 1). We estimated that because our validation cohort included approximately 80.0% of the population in England, approximately 13.9% of people in our data were part of the original cohort of 6 million patients used to develop the QCovid model.

Table 1 shows the baseline characteristics of patients. Of all patients, 16 599 875 (47.57%) were men and 6 052 563 (17.34%) were of ethnic minority background. The mean age was 51.1 years, which was slightly higher than in the cohort used to derive the QCovid models (48.2 years). For most pre-existing conditions, the estimated prevalence in the ONS Public Health Linked Data Asset is similar to the prevalence in the QResearch derivation cohort. However, because the ONS dataset is based on primary care data that did not contain a

list of codes as detailed as in the data used to develop the algorithm, the proportion of people taking anti-leukotriene or long-acting β_2 agonists or being prescribed oral steroids in the past 6 months was somewhat higher in our data than in the cohort used to derive the QCovid models.

26 985 (0.08%) patients had a COVID-19-related death during the first period (Jan 24 to April 30, 2020). 13 177 (0.04%) patients had a COVID-19-related death during the second period (May 1 to July 28, 2020). Of the 49 461 COVID-19 deaths that occurred in England over the study period, 40 162 (81.2%) were included in our data (appendix p 1). Coverage was lowest in London (6359 [74.20%] of 8570) and highest in the north west (6700 [84.6%] of 7923). In both periods, COVID-19 deaths occurred across all regions, with the greatest numbers in London in the first period (5403; 20.02% of all deaths) and in the north west in the second period (2411; 18.30% table 1). Of those who died in the first period, 15 334 (56.82%) were men, 11 651 (43.18%) were women, 4523 (16.76%) were from ethnic minority groups, 22 538 (83.52%) were aged 70 years and older, 8700 (32.24%) had diabetes, 8293 (30.73%) had dementia, and 6990 (25.90%) were identified as living in a care home (table 1). Those who had a COVID-19-related death in the second period had a similar profile to those in the first period but were older (11 647 [88.4%] aged 70 years and older) and more likely to live in a care home (4138 [31.40%]).

Table 2 shows the performance of the risk equations in the validation cohort for women and men in the two time periods. Overall, the values for the r^2 , D statistics, and C statistics were high and similar in women and men in both periods (table 2). In the first period, the equation explained 76.3% (95% CI 76.0–76.6) of the variation in time to COVID-19 death for women and 77.1% (76.9–77.4) for men (table 2). All these discrimination metrics were higher than in the original QResearch cohort used to validate the algorithm. The results were similar for the second validation period (table 2). Similar results were obtained when restricting the sample to 14 104 452 patients registered with practices using the TPP system (appendix p 1). Metrics obtained when restricting the sample to patients with valid BMI information were similar but marginally lower than

For the code see https://github.com/ONShmh/QCovid_validation

See Online for appendix

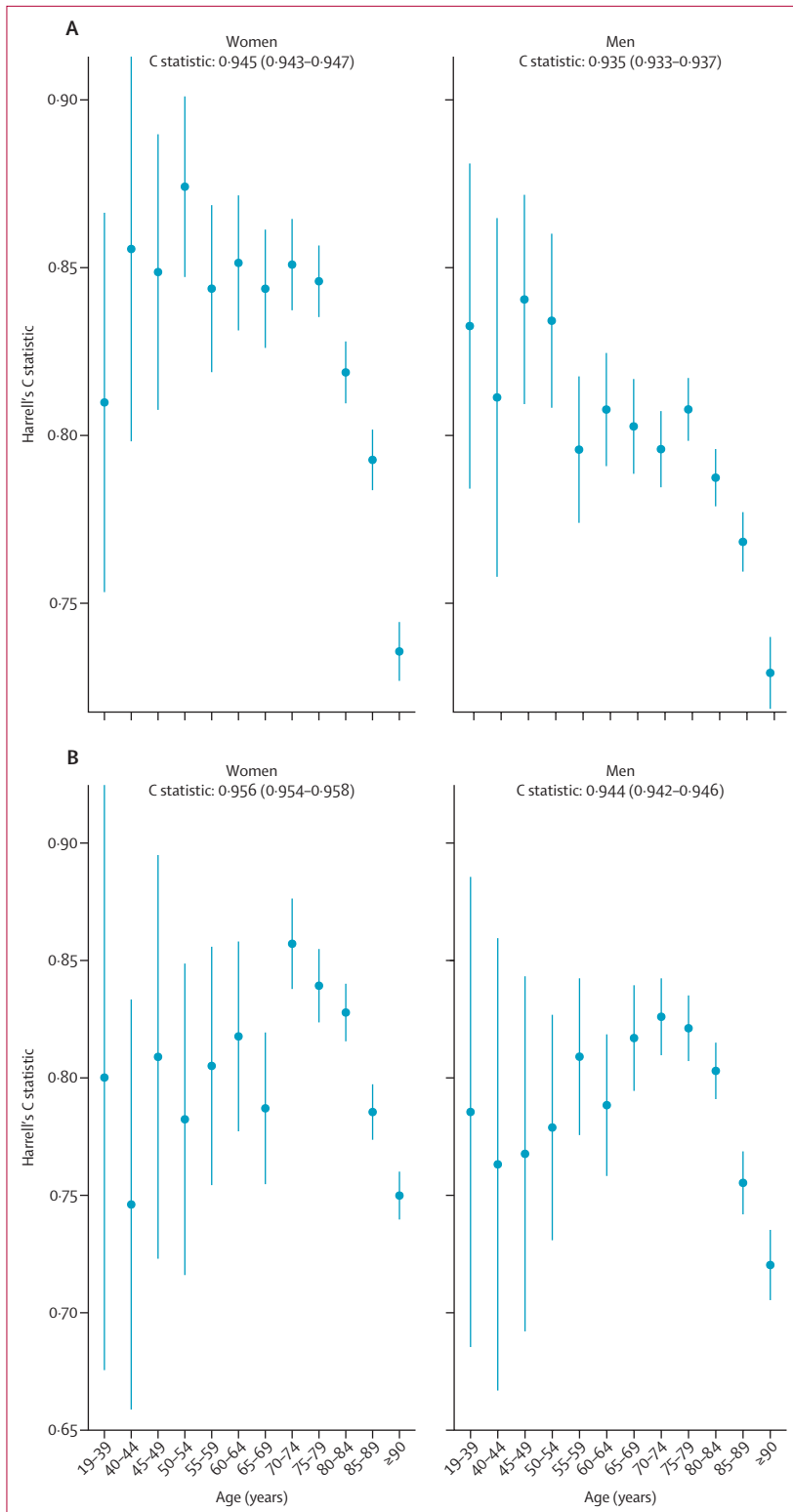


Figure 1: Harrell's C statistic by age group for men and women in the first period and second period
 (A) Results for the first period (Jan 24, to April 30, 2020). (B) Results for the second period (May 1, to July 28, 2020).
 Bars represent 95% CIs.

those obtained with the full sample (appendix p 2). Metrics for an alternative second period (May 1, to June 30, 2020; the period that was used in the study that developed the algorithm¹⁴) were similar (appendix p 2).

Figure 1 displays Harrell's C statistic by age group for men and women in the first period and second period. The Harrell's C statistics were greater than 0.700 for all age bands, indicating that even within each age band the model discriminates well (figure 1). The C statistics were lower for patients aged 90 years or older than for younger patients. The C statistic, r^2 , D statistic, and Brier score by age group, deprivation quintile, and ethnic group in men and women for both periods are reported in the appendix (pp 2–9). Performance was generally similar to that in the overall population, except for age where the performance was lower within individual age groups compared with in the overall population (appendix pp 2–9).

Figure 2 displays the calibration plots for the COVID-19 mortality equation for men and women and in the first period (this analysis was not done for the second period). Overall, both sets of equations were well calibrated because the predicted and observed risks were similar (figure 2). However, as in the original QResearch validation cohort, the model underestimated the risk of COVID-19 death for those in the top 5% of the predicted risk score (figure 2). We obtained similar results when restricting the sample to patients registered with practices using the TPP system (appendix p 12).

Figure 3 shows the sensitivity values for the mortality equation in the first period and second period assessed at different thresholds on the basis of the centiles of the predicted absolute risk in the validation cohort. Full results are reported in the appendix (p 10). Sensitivity was higher in women than in men and in the second period than in the first period (figure 3). In the first period, 65.94% of deaths in men occurred in those in the top 5% for predicted absolute risk of death from COVID-19 (90-day predicted absolute risks greater than 0.29%) and 71.67% of deaths in women occurred in the top 5% (predicted absolute risks greater than 0.19%; figure 3). In the second period, 71.10% of deaths occurred in men in the top 5% for predicted absolute risk of death from COVID-19 (predicted absolute risks greater than 0.278%) and 77.16% of deaths occurred in women in the top 5% (predicted absolute risks greater than 0.181%). Sensitivity for the two time periods based on relative risks is shown in the appendix (p 12; defined as the ratio of the individual's predicted absolute risk to the predicted absolute risk for a person of the same age and sex with a White ethnicity, BMI 25 kg/m², and mean deprivation score with no other risk factors). In the first period, 40.56% of deaths occurred in men and 42.63% in women in the top 5% for predicted relative risk of death from COVID-19 (figure 2; appendix p 13). In the second period, 42.62% of deaths occurred in men and 43.57% in

women in the top 5% for predicted relative risk of death from COVID-19.

We report the distribution of predicted risks of COVID-19 death by age group and sex in the appendix (p 14). The predicted risk increased exponentially with age and we found substantial variation in predicted risks within age group (appendix p 14).

Discussion

We validated the QCovid clinical risk prediction model for mortality due to COVID-19 using a national external linked dataset. We used national linked datasets from the 2011 census, GP, and death registry data for a population-representative sample of nearly 35 million adults. The risk models had excellent discrimination, were well calibrated (predicted and observed risks were similar) and had a high sensitivity (two-thirds or more of deaths occurred in the people in the top 5% for predicted absolute risk of death from COVID-19).

Our study had several important strengths. First, we used a unique linked dataset based on the 2011 census for nearly 35 million people living in England. Second, we used various metrics over two time periods to validate the QCovid predictive model. All the performance metrics in the two time periods for both men and women indicated that the algorithm performs well, despite the demographic profile of people who died being slightly different in the two periods. The metrics were similar to those of the original validation of QCovid in the QResearch database.¹⁴ The model performance was even slightly higher than in the derivation cohort, probably because of broader variation in risk factors in this larger cohort. Finally, we showed that the model's performance was similar when restricting the sample to patients who were registered with practices using a different clinical computer system provider (TPP) and therefore not used to derive the QCovid model.

This study also has several limitations. First, because of data limitations, we could not derive all predictors in the same way as in the derivation cohort. Despite these inconsistencies, the model had excellent discrimination and calibration. Second, we only focused on COVID-19-related deaths and not hospital admissions because of the lack of data. Additionally, early in the pandemic some COVID-19-related deaths might not have been recorded as such. Third, our sample only contained data from approximately 80% of the population aged 19–100 years in England. Because the Public Health Data Asset was based on the 2011 census, the data excluded approximately 6% of people who lived in England in 2011 but did not take part in the 2011 census. Additionally, the data also excluded approximately 5% of 2011 census respondents who could not be linked to the 2011–13 NHS patients register. Because the dataset was based on individuals enumerated at the 2011 census, people who had immigrated to England since 2011 were excluded. However, recent migrants tend to be younger than the

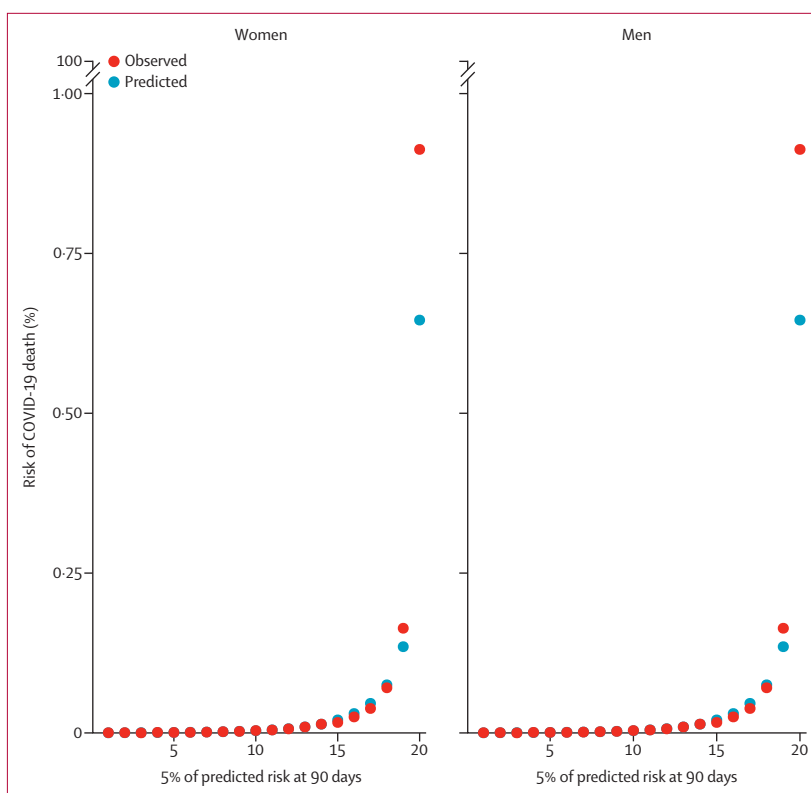


Figure 2: Predicted and observed risk of COVID-19-related death in the first study period
First study period was Jan 24, to April 30, 2020.

native population and therefore at lower risk of COVID-19 death.²⁵ Our data also excluded people who were not registered with the NHS. Another limitation is that an estimated 13·9% of patients in our cohort were also part of the cohort used for deriving the QCovid model. However, we found that the model's performance was similar when using only a subset of patients who were registered with practices using TPP and, therefore, were not part of the model development cohort, which used data from patients registered with EMIS practices. Comparing the performance of QCovid to that of other risk prediction models in the ONS Public Health Data Asset is an important area for further research.

QCovid represents a new approach for population risk stratification for adverse outcomes from COVID-19 and our validation indicates that the risk algorithm performs well on external data not used for the algorithm's derivation. A companion study that is currently underway is aiming to externally validate additional QCovid algorithms that use datasets from Wales (SAIL¹⁶) and Scotland (EAVE-II¹⁷), the results of which are to be reported separately. Moreover, despite the QCovid algorithm discussed here being specifically designed to inform UK health policy and interventions to manage COVID-19-related risks, the algorithm also has international potential, subject to local validation. The

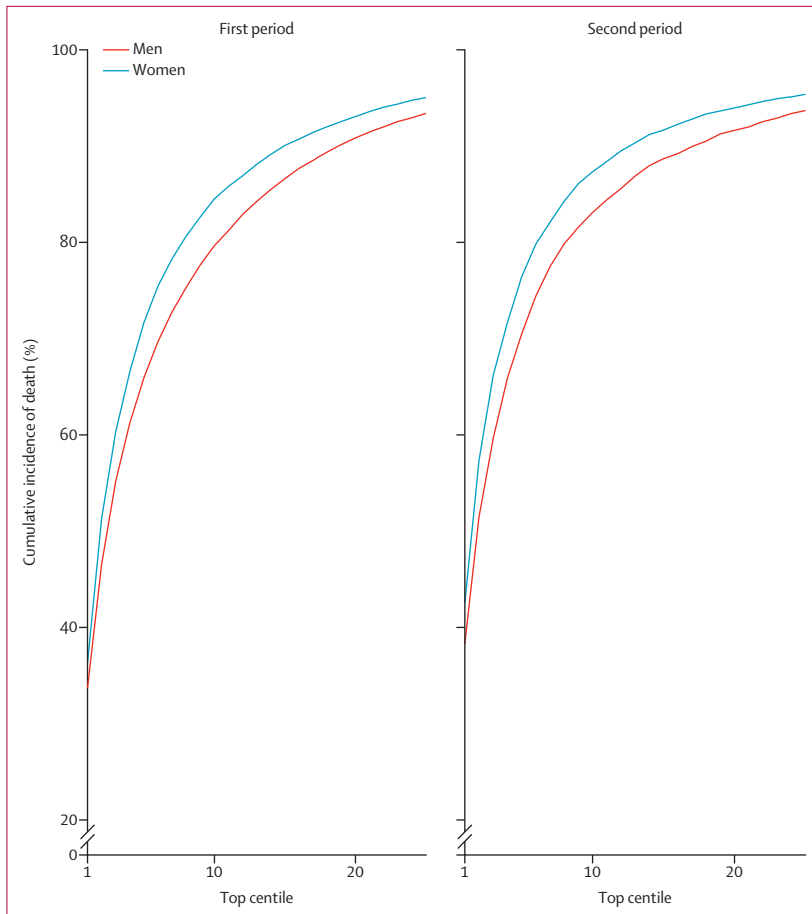


Figure 3: Sensitivity for COVID-19-related death in the validation cohort for the first and second study periods
 The first study period was Jan 24, to April 30, 2020, and the second study period was May 1, to July 28, 2020. Centiles were based on predicted absolute risks in men and women in each period. Sensitivity (cumulative percentage of deaths) is percentage of total deaths in the period that occurred within the group of patients above the predicted risk threshold.

QCovid risk model predicts COVID-19 deaths in the general population over a fairly short time period, which potentially limits the algorithm's applicability. Predictive models that operate over longer time periods are needed. QCovid could nonetheless be deployed in several health and care applications, either during the current phase of the pandemic, or in subsequent waves of infection. These applications could include supporting targeted recruitment for clinical trials, vaccine prioritisation, and discussions between patients and clinicians in relation to work and health risks, for example through weight reduction given that obesity is the single most important modifiable risk factor for serious COVID-19 complications.⁷ However, using the model to allow additional exposure for people with low predicted risk would warrant additional analysis and close monitoring of the consequences.

In conclusion, this study presents a robust validation of a new prediction model that could be used to support population risk stratification in relation to public health

interventions, for example vaccine use. We anticipate that the algorithms will be updated regularly as understanding of COVID-19 increases, more data become available, new variants emerge, effective treatments for COVID-19 become available, the vaccination programme rolls out, immunity levels change, or as behaviour in the population changes (eg, reduced adherence to physical distancing rules) and hence we anticipate that this validation will need to be repeated on a regular basis. The existence of a common appropriately developed model that is evidence based, consistently implemented, and supported by the academic, clinical, and patient communities is important for patients, carers, and clinicians. Use of this model will then help ensure consistent policy and clear national communication between policy makers, professionals, employers, and the public.

Contributors

Study conceptualisation was led by NM, JH-C, VN, and CC. All authors contributed to the development of the research question and study design, with development of advanced statistical aspects led by CC and VN. VN, RS, PB, PP, JH-C, and JM were involved in data specification, curation, and collection. JH-C developed, checked, or updated clinical code groups. VN led the statistical analyses, which were checked by LL. All authors contributed to the interpretation of the results. VN and JH-C wrote the first draft of the paper. All authors contributed to the critical revision of the manuscript for important intellectual content and approved the final version of the manuscript. All authors had full access to all the data in the study and had final responsibility to submit for publication. The lead author (VN) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned have been explained.

Declaration of interests

JH-C reports grants from the National Institute for Health Research Biomedical Research Centre, Oxford, UK; John Fell Oxford University Press Research Fund; Cancer Research UK, through the Cancer Research UK Oxford Centre; and the Oxford Wellcome Institutional Strategic Support Fund, during the conduct of the study. JH-C is an unpaid director of QResearch, a not-for-profit organisation that is a partnership between the University of Oxford, Oxford, UK, and EMIS Health, who supplied the QResearch database used for this work. JH-C is a founder and shareholder of ClinRisk and was the company's medical director until May 31, 2019. ClinRisk produces open and closed source software to implement clinical risk algorithms (outside this work) into clinical computer systems. All other authors declare no competing interests.

Data sharing

The ONS Public Health Linked Data Asset will be made available on the ONS Secure Research Service for accredited researchers. Researchers can apply for accreditation through the ONS Research Accreditation Service (<https://www.ons.gov.uk/aboutus/whatwedo/statistics/requestingstatistics/approvedresearcherscheme>). The data will include all variables used in this analysis, except predictors that are based on radiotherapy and systemic chemotherapy records, which cannot be shared.

Acknowledgments

We acknowledge the contribution of Jane Campbell, Nazmus Haq, Joanna Moody, and Shamim Rahman from the UK Department of Health and Social Care; and Joy Preece and Dan Ayoubkhani from the ONS. This project involved data derived from patient-level information collected by the NHS, as part of the care and support of patients with cancer. Access to the data was facilitated by the Public Health England Office for Data Release. The Hospital Episode Statistics data used in this analysis are re-used by permission from NHS Digital who retain the copyright for these data.

References

- 1 Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020; **395**: 1054–62.
- 2 Yancy CW. COVID-19 and African Americans. *JAMA* 2020; **323**: 1891–92.
- 3 Chen T, Wu D, Chen H, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. *BMJ* 2020; **368**: m1091-m.
- 4 Weiss P, Murdoch DR. Clinical course and mortality risk of severe COVID-19. *Lancet* 2020; **395**: 1014–15.
- 5 Wadhwa RK, Wadhwa P, Gaba P, et al. Variation in COVID-19 hospitalizations and deaths across New York City boroughs. *JAMA* 2020; **323**: 2192–95.
- 6 Le Brocq S, Clare K, Bryant M, Roberts K, Tahrani AA. Obesity and COVID-19: a call for action from people living with obesity. *Lancet Diabetes Endocrinol* 2020; **8**: 652–54.
- 7 Sattar N, McInnes IB, McMurray JJV. Obesity is a risk factor for severe COVID-19 infection: multiple potential mechanisms. *Circulation* 2020; **142**: 4–6.
- 8 Singh AK, Gillies CL, Singh R, et al. Prevalence of co-morbidities and their association with mortality in patients with COVID-19: a systematic review and meta-analysis. *Diabetes Obes Metab* 2020; **22**: 1915–24.
- 9 Smith GD, Spiegelhalter D. Shielding from covid-19 should be stratified by risk. *BMJ* 2020; **369**: m2063
- 10 Wynants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19 infection: systematic review and critical appraisal. *BMJ* 2020; **369**: m1328.
- 11 Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. *BMJ* 2017; **357**: j2099.
- 12 Hippisley-Cox J, Coupland C. Development and validation of QMortality risk prediction algorithm to estimate short term risk of death and assess frailty: cohort study. *BMJ* 2017; **358**: j4208.
- 13 Hippisley-Cox J, Coupland C. Development and validation of QDiabetes-2018 risk prediction algorithm to estimate future risk of type 2 diabetes: cohort study. *BMJ* 2017; **359**: j5019.
- 14 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.
- 15 Williamson EJ, Walker AJ, Bhaskaran K, et al. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020; **584**: 430–36
- 16 Hollinghurst J, Lyons J, Fry R, et al. The impact of COVID-19 on adjusted mortality risk in care homes for older adults in Wales, United Kingdom: a retrospective population-based cohort study for mortality in 2016–2020. *Age Ageing* 2021; **50**: 25–31.
- 17 Simpson CR, Robertson C, Vasileiou E, et al. Early pandemic evaluation and enhanced surveillance of COVID-19 (EAVE II): protocol for an observational study using linked Scottish national data. *BMJ Open* 2020; **10**: e039097.
- 18 Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015; **162**: 55–63.
- 19 Benchimol EI, Smeeth L, Guttman A, et al. The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS Med* 2015; **12**: e1001885.
- 20 Hippisley-Cox J, Clift AK, Coupland CAC, et al. Protocol for the development and evaluation of a tool for predicting risk of short-term adverse outcomes due to COVID-19 in the general UK population. *medRxiv* 2020; published online June 29. <https://doi.org/10.1101/2020.06.28.20141986> (preprint).
- 21 Royston P. Explained variation for survival models. *Stata J* 2006; **6**: 1–14.
- 22 Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. *Stat Med* 2004; **23**: 723–48.
- 23 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; **15**: 361–87.
- 24 Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiol* 2010; **21**: 128–38.
- 25 Rienzo C, Vargas-Silva C. Migrants in the UK: an overview. COMPAS, University of Oxford. Nov 6, 2020. <https://migrationobservatory.ox.ac.uk/wp-content/uploads/2017/02/Briefing-Migrants-in-the-UK-An-Overview.pdf> (accessed May 18, 2021).