

Remote COVID-19 Assessment in Primary Care (RECAP) risk prediction tool: derivation and real-world validation studies



Ana Espinosa-Gonzalez*, Denys Prociuk*, Francesca Fiorentino, Christian Ramtale, Ella Mi, Emma Mi, Ben Glampson, Ana Luisa Neves, Cecilia Okusi, Laiba Husain, Jack Macartney, Martina Brown, Ben Browne, Caroline Warren, Rachna Chowla, Jonty Heaversedge, Trisha Greenhalgh, Simon de Lusignan, Erik Mayer, Brendan C Delaney

Summary

Background Accurate assessment of COVID-19 severity in the community is essential for patient care and requires COVID-19-specific risk prediction scores adequately validated in a community setting. Following a qualitative phase to identify signs, symptoms, and risk factors, we aimed to develop and validate two COVID-19-specific risk prediction scores. Remote COVID-19 Assessment in Primary Care-General Practice score (RECAP-GP; without peripheral oxygen saturation [SpO₂]) and RECAP-oxygen saturation score (RECAP-O2; with SpO₂).

Methods RECAP was a prospective cohort study that used multivariable logistic regression. Data on signs and symptoms (predictors) of disease were collected from community-based patients with suspected COVID-19 via primary care electronic health records and linked with secondary data on hospital admission (outcome) within 28 days of symptom onset. Data sources for RECAP-GP were Oxford-Royal College of General Practitioners Research and Surveillance Centre (RCGP-RSC) primary care practices (development set), northwest London primary care practices (validation set), and the NHS COVID-19 Clinical Assessment Service (CCAS; validation set). The data source for RECAP-O2 was the Doctaly Assist platform (development set and validation set in subsequent sample). The two probabilistic risk prediction models were built by backwards elimination using the development sets and validated by application to the validation datasets. Estimated sample size per model, including the development and validation sets was 2880 people.

Findings Data were available from 8311 individuals. Observations, such as SpO₂, were mostly missing in the northwest London, RCGP-RSC, and CCAS data; however, SpO₂ was available for 1364 (70.0%) of 1948 patients who used Doctaly. In the final predictive models, RECAP-GP (n=1863) included sex (male and female), age (years), degree of breathlessness (three point scale), temperature symptoms (two point scale), and presence of hypertension (yes or no); the area under the curve was 0.80 (95% CI 0.76–0.85) and on validation the negative predictive value of a low risk designation was 99% (95% CI 98.1–99.2; 1435 of 1453). RECAP-O2 included age (years), degree of breathlessness (two point scale), fatigue (two point scale), and SpO₂ at rest (as a percentage); the area under the curve was 0.84 (0.78–0.90) and on validation the negative predictive value of low risk designation was 99% (95% CI 98.9–99.7; 1176 of 1183).

Interpretation Both RECAP models are valid tools to assess COVID-19 patients in the community. RECAP-GP can be used initially, without need for observations, to identify patients who require monitoring. If the patient is monitored and SpO₂ is available, RECAP-O2 is useful to assess the need for treatment escalation.

Funding Community Jameel and the Imperial College President's Excellence Fund, the Economic and Social Research Council, UK Research and Innovation, and Health Data Research UK.

Copyright © 2022 The Author(s). Published by Elsevier Ltd. This is an Open Access article under the CC BY 4.0 license.

Introduction

Since the start of the COVID-19 pandemic in 2019, a priority in health systems globally has been to rapidly develop and validate data-driven algorithms to predict risk and guide care.^{1–4} This includes a requirement to accurately distinguish between patients with COVID-19 who can be safely managed in the community setting and those whose care should be escalated to hospital. In primary care, patients with COVID-19 have been typically categorised into three care strategy categories: reassure and safety net (green risk), monitor (amber risk), and admit to hospital

(red risk),⁵ on the basis of clinical judgement and consensus using a few clinical parameters, such as heart rate, pulse rate, and peripheral oxygen saturation (SpO₂).

Several risk prediction models have been used to support the management of COVID-19. However, these were derived and validated in hospital in-patient populations.^{6,7} The QCOVID score, based on primary care data, predicts risk of death from COVID-19 based on demographic factors and pre-existing medical conditions, but it does not predict acute deterioration based on current clinical observations.⁸ The National Early Warning Score 2

Lancet Digit Health 2022

Published Online
July 28, 2022
[https://doi.org/10.1016/S2589-7500\(22\)00123-6](https://doi.org/10.1016/S2589-7500(22)00123-6)
See Online/Comment
[https://doi.org/10.1016/S2589-7500\(22\)00146-7](https://doi.org/10.1016/S2589-7500(22)00146-7)

*Contributed equally

Institute of Global Health Innovation, Department of Surgery and Cancer, Imperial College London, London, UK (A Espinosa-Gonzalez PhD, D Prociuk MSc, F Fiorentino PhD, C Ramtale MSc, El Mi MBBChir, Em Mi MBBChir, A L Neves PhD, E Mayer PhD, Prof B C Delaney MD); Department of Surgery and Cancer, Imperial College Healthcare NHS Trust, London, UK (B Glampson); Nuffield Department of Primary Care, University of Oxford, Oxford, UK (C Okusi MRes, L Husain, J Macartney, Prof T Greenhalgh PhD, Prof S de Lusignan MD); South Central Ambulance Service NHS Trust, Otterbourne, UK (M Brown MSc, B Browne BM, C Warren BM); King's Health Partners, London, UK (R Chowla MBA); South East London NHS Clinical Commissioning Group, London, UK (J Heaversedge MBBS); Nightingale-Saunders Clinical Trials & Epidemiology Unit, King's Clinical Trials Unit, King's College London, London, UK (F Fiorentino)

Correspondence to:
Prof Brendan C Delaney, Institute of Global Health Innovation, Department of Surgery and Cancer, Imperial College London, London W2 1NY, UK
brendan.delaney@imperial.ac.uk

Research in context

Evidence before this study

This study was conceived in March and April 2020, during the first COVID-19 peak in the UK. The review was done according to Cochrane Collaboration standards for rapid reviews; we identified systematic reviews and large-scale observational studies describing the signs and symptoms of COVID-19. We searched for papers published between Jan 1, 2020, and March 31, 2020, and repeated for papers published until April 30, 2020. We searched PubMed and Embase for research on COVID-19, using the search terms "COVID-19", "SARS-CoV-2", "novel corona", and "2019-ncov", "diagnos\$", "prognos\$", and "prediction model". We also searched the publicly available publication list of the COVID-19 living systematic review using the same terms. Evidence gathered showed worsening of COVID-19 symptoms around the seventh day of disease and challenges to identify patients with higher likelihood of severity to emphasise their monitoring. To this end, tools such the National Early Warning Score 2 have been used in the UK to assess patients with COVID-19 in primary care, but they do not capture the characteristics of COVID-19 infection and are not suitable for community remote assessment. Several COVID-19 risk scores have been developed. QCOVID provides a risk of mortality considering patients' existing risk factors but does not include acute signs and symptoms. The International Severe Acute Respiratory and emerging Infection Consortium 4C Deterioration model was specifically developed for hospital settings. In England, the NHS has implemented the Oximetry @home strategy to monitor patients with acute COVID-19 deemed at risk (>64 years old or with comorbidities) by providing pulse oximeters; however, the

criteria for monitoring or for escalation have not been validated. There was a need to develop a risk prediction score to establish COVID-19 patients' risk of deterioration to be used in the community via both face-to-face or remote consultation.

Added value of this study

We developed and validated two COVID-19-specific risk prediction scores: one to be used in the initial remote assessment of patients with COVID-19 to assess need for monitoring (Remote COVID-19 Assessment in Primary Care-General Practice score; RECAP-GP) and another to assess the need for treatment escalation, which includes peripheral saturation of oxygen as one of the model predictors (RECAP-oxygen saturation score; RECAP-O2). To our knowledge, these scores are the first COVID-19-specific risk prediction scores to assess and monitor COVID-19 patients' risk of deterioration remotely. These scores will be a valuable resource to complement the use of oximetry in the community when assessing a patient with COVID-19.

Implications of all the available evidence

To manage pandemic waves and their demand on health care, patients with COVID-19 require close monitoring in the community and prompt escalation of their treatment. Guidance available so far relied on unvalidated tools and clinician judgement to assess deterioration. COVID-19-specific community-based risk prediction scores, such as RECAP, might contribute to reducing the uncertainty in the assessment and monitoring of COVID-19 patients, increase safety in clinical practice, and improve outcomes by facilitating rapid treatment escalation.

(NEWS2) score is widely used in UK emergency care as an early warning score for sepsis. However, the parameters in NEWS2 (such as tachycardia, fever, and hypotension) are usually very late signs of clinical deterioration, and they have been found to perform poorly in the acute assessment of suspected COVID-19 in hospital in-patients,⁹ and have not been evaluated outside hospital settings.¹⁰ Because initial patient contacts with health services are increasingly carried out remotely, a score was needed that could be administered over the telephone or other remote means.¹¹ Instead of vital signs, a score could be based on clinical symptoms which the patient or a relative could assess (eg, perceived breathlessness or confusion). A Delphi study (with qualitative and survey components) using 112 primary care clinicians and 50 patients derived a set of data items comprising symptoms and vital signs that could be included in a putative remote COVID Assessment in Primary Care prediction (RECAP) model.¹² Templates for collection of these RECAP data elements (known as RECAP-V0), using appropriate Systematized Nomenclature of Medicine (SNOMED) clinical terms, were developed for electronic

health record systems so that a model could be derived and then validated.

This study aimed to develop and validate two prediction models: (1) a score incorporating observable vital signs (heart rate, temperature, respiratory rate, and oxygen saturation) and (2) a score for use when these parameters cannot be measured due to an absence of equipment or poor patient familiarity with self-assessment in a virtual consultation. Two cutoff values were derived and validated: one for the need of monitoring (green and amber risk) and the other for hospital admission (amber and red risk) of patients with COVID-19.

Methods

Study design and participants

RECAP was a prospective cohort and observational study. Patients (≥ 18 years old) with symptoms of COVID-19 who presented to primary care within 14 days of onset of symptoms diagnosed on the basis of clinical judgement who were able to provide informed consent (except for people who used Doctaly Assist where consent was not sought) were eligible for inclusion. Patients were

followed for 28 days from the onset of symptoms to determine the occurrence of hospital admissions due to COVID-19. The study protocol¹³ and statistical analysis plan¹⁴ have been published previously.

The study was sponsored by Imperial College London and approved by the North West - Greater Manchester East Research Ethics Committee and Health Research Authority in May 2020 (Integrated Research Assessment System [IRAS] number 283024; North West - Greater Manchester East Research Ethics Committee and Health Research Authority Research Ethics Committee reference number 20/NW/0266). The study was badged as an Urgent Public Health Study by the National Institute of Health Research in October, 2020. The Northwest London Whole System Integrated Care data analysis was undertaken within a research database that was given favourable ethics approval by the West Midlands Solihull Research Ethics Committee (reference number 18/WM/0323; IRAS project ID 252449). All data used in this paper were fully anonymised before analysis. Imperial clinical analytics, research, and evaluation environment (iCARE) is a Trusted Research Environment (TRE) and provides access to Health Research Authority Research Ethics Committee-approved anonymised data for research (reference number 21/SW/0120; IRAS project ID 282093). Analysis was done within the secure data processing platform of the Oxford Royal College of General Practitioners Clinical Informatics Digital Hub (ORCHID) TRE. Access to Doctaly Assist data was granted under the Control of Patient Information (COPI) regulation (and Confidentiality Advisory Group Resolution 5 after expiration of COPI in March, 2022); therefore, patient consent was not required. However, patients could opt-out using the NHS Digital National Opt-Out policy.

Data

To allow for parallel derivation and validation of the RECAP scores on different cohorts, we used four UK primary care settings (figure 1). General practices (GPs) from the northwest London and clinical research network and practices belonging to the Oxford–Royal College of General Practitioners (RCGP) Research and Surveillance Centre (RSC) network,^{15,16} which includes practices from all over England, collected data from Oct 1, 2020, to Feb 28, 2021. The COVID-19 Clinical Assessment Service (CCAS), a nationwide COVID-19 remote assessment service, collected patient data from March 15, 2021, to May 23, 2021. We also used data collected from Nov 26, 2020, to Oct 26, 2021 through the Doctaly Assist platform, which was used in southeast London (the main contributor to which was the Lewisham clinical commissioning group) for monitoring patients with COVID-19.¹⁷ We ensured that datasets used for development and validation of each model did not overlap geographically or in time (figure 1). We excluded northwest London practices from the RSC cohort.

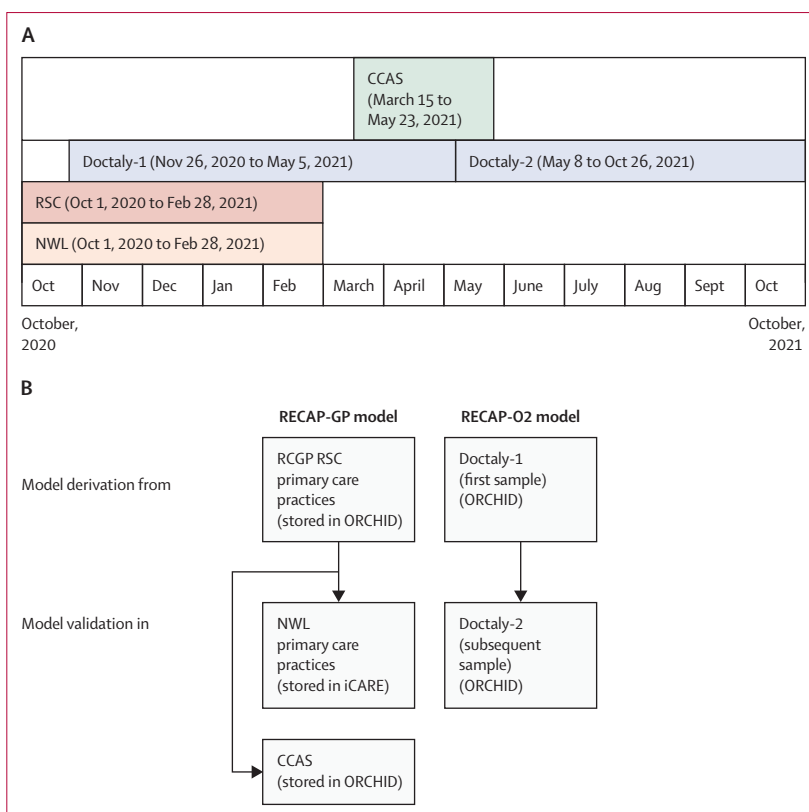


Figure 1: Settings used for derivation and validation of the RECAP scores.

(A) Recruitment dates per data source shown. Geographical separation as follows: CCAS was accessible for all patients in England; Doctaly was accessible for patients in southeast London; the RSC network was used by practices all over England (practices in northwest London were excluded), northwest London practices are restricted to northwest London boroughs. (B) Data sources used for derivation and validation of the RECAP scores. CCAS=NHS-111 COVID clinical assessment service. iCARE=Imperial clinical analytics, research, and evaluation environment. NWL=northwest London. ORCHID=Oxford-Royal College of General Practitioners Clinical Informatics Digital Hub environment. RCGP RSC=Royal College of General Practitioners Research and Surveillance Centre. RECAP-GP=Remote COVID-19 Assessment in Primary Care-General Practice score. RECAP-O2=Remote COVID-19 Assessment in Primary Care-oxygen saturation score.

Patients in southeast London were directly seen in COVID-19 hot hubs during the period of study and onboarded to Doctaly.

We used primary data on patients' signs and symptoms collected in the community at the point of consultation linked to secondary data on hospital outcomes. For data collection, the RECAP-V0 electronic template, with selected SNOMED codes,¹² was used in electronic health record systems (EMIS, TPP SystemOne, and Adastra) and completed by clinicians when assessing patients with signs and symptoms of COVID-19. For the Doctaly Assist platform, patients were provided with home pulse oximeters and onboarded to a platform that uses WhatsApp¹⁸ to ask questions based on RECAP-V0 items.^{17,19,20} Additional data on comorbidities, ethnicity (subsequently grouped into two categories White and minority ethnic [or non-White]), age, and sex that were included in the development of the risk prediction model were available from the electronic health record. Since the clinical

	RCGP RSC (n=1863)	NWL (n=2415)	CCAS (n=2674)	Doctaly-1 (n=1948)	Doctaly-2 (n=2085)
Age, years	49 (17.7)	46 (18.3)	42 (15.9)	44 (13.1)	39 (11.9)
BMI					
BMI, kg/m ²	28.8 (6.6)	26 (6.5)	28.8 (7.1)	NA	NA
Missing data	124 (6.6%)	84 (3.5%)	352 (13.1%)	NA	NA
Obesity (BMI ≥30 kg/m ²)	633 (34.0%)	545 (22.6%)	970 (36.3%)	NA	NA
Missing data	124 (6.6%)	84 (3.5%)	358 (13.4%)	NA	NA
Sex					
Female	1061 (56.9%)	1352 (56.0%)	1470 (55.0%)	1246 (64.0%)	1271 (61.0%)
Male	802 (43.1%)	1063 (44.0%)	1204 (45.0%)	702 (36.0%)	813 (39.0%)
Ethnicity					
Non-White	624 (33.5%)	1787 (74.0%)	2105 (78.7%)	1334 (68.5%)	1063 (51.0%)
White	1239 (66.5%)	628 (26.0%)	569 (21.3%)	614 (31.5%)	1022 (49.0%)
Missing data	121 (6.5%)	207 (8.6%)	438 (16.4%)	146 (7.5%)	2018 (96.8%)
Comorbidity					
Diabetes*	204 (11.0%)	369 (15.3%)	157 (5.9%)	NA	NA
Hypertension*	465 (25.0%)	565 (23.3%)	310 (11.6%)	NA	NA
Coronary heart disease*	130 (7.0%)	72 (3.0%)	101 (3.8%)	NA	NA
Chronic kidney disease*	130 (7.0%)	91 (3.8%)	53 (2.0%)	NA	NA
Adverse social circumstances*	3 (0.2%)	3 (0.2%)	NA	NA	NA

Data are n (%) or mean (SD). BMI=body-mass index. NA=not available. RCGP RSC=Royal College of General Practitioners Research and Surveillance Centre. NWL=northwest London. CCAS=COVID-19 Clinical Assessment Service. *Data obtained from linked general practitioner electronic health records.

Table 1: Summary population characteristics

See Online for appendix

question that needs to be supported is “does this patient require care escalation?”, hospital admission, defined as a night’s hospital stay, was the main outcome to be predicted by the model. To derive outcomes, data were linked with hospital episode statistics via the ORCHID TRE at Oxford University, or with the northwest London SitRep data on COVID-19 admissions in iCARE environment at Imperial College Healthcare NHS Trust. When identifying patients’ COVID-associated admissions in hospital episode statistics data, we searched for COVID-19 International Classification of Diseases-10 codes U071, U072, U073, or U074 as first, second, or third cause of admission. Patients with COVID-19 as the second or third cause of hospital administration were included in this study if the first cause of admission was pneumonia, dyspnoea, pulmonary embolism, or chest pain.²¹

Statistical analysis

We estimated a minimum sample size of 1317 participants for model development and 1400 for model external validation assuming 10% hospitalisation rate for COVID-19, a maximum of 24 predictor variables, a binary outcome (hospital admission), and a minimum 85% model specificity on validation. We aimed to recruit at least 2880 participants per setting assuming 5–6% loss

to follow-up.¹⁴ A study-specific SNOMED code carrying the RECAP National Institute of Health research portfolio number was used to identify the relevant records for the study in ORCHID and iCARE.

A detailed explanation of analysis is provided in the published statistical analysis plan¹⁴ and the appendix (pp 2–3). The extent of missing data for each variable (outcome and predictors) was assessed on degree of missingness and pattern of missingness (at random or not at random) through visual inspection of the distribution. Missing data were then imputed using multiple imputation chain equations, including the outcome endpoint. The imputed datasets were stacked using Rubin’s rule before analysis. A probabilistic risk prediction based on a multivariable logistic regression model, including the variables in RECAP-V0 as factors, was done for the RSC dataset (RECAP-GP), and separately for the Doctaly Assist dataset (RECAP-O2).

The second model that used only Doctaly Assist data was planned given the predicted availability of larger numbers of oxygen saturation readings in this dataset. The models allow estimation of the likelihood of a particular patient with a COVID-19 diagnosis being admitted to hospital with COVID-19 within 28 days of symptom onset. Variables were checked for independence from each other by including in the model interaction terms between age and respiratory rate when available. Elements of the RECAP-V0 that were constructed using alternative codes representing different levels of severity had that relationship maintained in the models.

Using backward elimination, in which the least significant variable was eliminated and the model rerun for each iteration, two models including only the predictor factors that were shown to be statistically significant with a p value of less than 0.05 were obtained. Internal model validation was done using bootstrapping (500 repeats) and then the receiver operating characteristic (ROC) curves were plotted. The model performance in subpopulations in external validation datasets (CCAS, Doctaly 2, and northwest London) by age (≥65 years and <65 years) and by sex was investigated by comparing the diagnostic measures Akaike information criterion (AIC), optimism-corrected C-statistic, and visual inspection of the ROC curves. The effect on the area under the curve (AUC) of adding oxygen saturation at rest as a non-linear term in the RECAP-O2 model was investigated. Selection of upper and lower cutpoints for creating a red, amber, and green categorisation of patients was based on clinical consideration of risk and model performance. External validation of the RECAP-GP model was done using data separately from northwest London primary care practices and CCAS to verify the specificity, sensitivity, negative predictive value, and positive predictive value of the model predictions. The RECAP-O2 model was validated using a subsequent cohort of participants.¹⁴ Software analyses were done with R (version 4.1.2). The main

	RCGP RSC (RECAP-GP model; n=1863)		NWL (RECAP-GP external validation; n=2415)		CCAS (RECAP-GP external validation; n=2674)		Doctaly-1 (RECAP-O2 model; n=1948)		Doctaly-2 (RECAP-O2 external validation; n=2085)	
	Completeness	Considered in model	Completeness	Considered in model	Completeness	Considered in model	Completeness	Considered in model	Completeness	Considered in model
Shortness of breath*	1602 (86.0%)	Yes	2101 (87.0%)	Yes	2567 (96.0%)	Yes	1792 (92.0%)	Yes	2085 (100%)	Yes
Feeling feverish or shivers*	1788 (96.0%)	Yes	1883 (78.0%)	Yes	2567 (96.0%)	Yes	1782 (91.5%)	Yes	2085 (100%)	Yes
Temperature (observed)	447 (24.0%)	No	169 (7.0%)	No	721 (27.0%)	No	837 (43.0%)	No	938 (45.0%)	No
Fatigue*	1713 (91.9%)	Yes	2101 (87.0%)	Yes	2433 (91.0%)	Yes	1772 (91.0%)	Yes	2085 (100%)	Yes
Acute cognitive decline*	1527 (82.0%)	Yes	1690 (70.0%)	Yes	0	No	1772 (91.0%)	Yes	2085 (100%)	Yes
Time from first symptoms (days)	1732 (93.0%)	No	772 (32.0%)	No	2567 (96.0%)	No	1928 (99.0%)	Yes	2085 (100%)	Yes
Respiratory rate	279 (15.0%)	No	169 (7.0%)	No	588 (22.0%)	No	1402 (72.0%)	Yes	1855 (89.0%)	Yes
Heart rate	409 (21.9%)	No	507 (21.0%)	No	1123 (42.0%)	No	1363 (70.0%)	No	20 (1.0%)	No
Oxygen saturation at rest	279 (15.0%)	No	386 (16.0%)	No	347 (13.0%)	No	1363 (70.0%)	Yes	1292 (62.0%)	Yes
Oxygen saturation after 40 steps	111 (6.0%)	No	193 (8.0%)	No	80 (3.0%)	No	1285 (66.0%)	Yes	1188 (57.0%)	Yes
Muscle aches	894 (48.0%)	No	941 (39.0%)	No	1390 (52.0%)	No	NA	No	NA	No
Trajectory of breathlessness*	838 (45.0%)	No	1267 (52.5%)	No	2112 (79.0%)	No	19 (1.0%)	yes	0	Yes
Diabetes	1863 (100%)†	Yes	2415 (100%)†	Yes	2674 (100%)‡	Yes	NA	No	NA	No
Hypertension	1863 (100%)†	Yes	2415 (100%)†	Yes	2674 (100%)‡	Yes	NA	No	NA	No
Coronary health disease	1863(100%)†	Yes	2415 (100%)†	Yes	2674 (100%)‡	Yes	NA	No	NA	No
Chronic kidney disease	1863 (100%)†	Yes	2415 (100%)†	Yes	2674 (100%)‡	Yes	NA	No	NA	No
Age	1863 (100%)†	Yes	2415 (100%)†	Yes	2674 (100.0%)‡	Yes	1948 (100.0%)	Yes	2085 (100%)	Yes
Sex	1863 (100%)†	Yes	2415 (100%)†	Yes	2674 (100%)‡	Yes	1948 (100%)	Yes	2085 (100%)	Yes
Ethnicity	1741 (93.5%)†	Yes	2197 (91.0%)†	Yes	2246 (84.0%)	Yes	1801 (92.5%)	Yes	62 (3.0%)	No

CCAS=COVID-19 Clinical Assessment Service. NA=not available. NWL=Nnorthwest London. RECAP-GP=Remote COVID-19 Assessment in Primary Care-General Practice score. RECAP-O2=Remote COVID-19 Assessment in Primary Care-oxygen saturation score. RCGP RSC=Royal College of General Practitioners Research and Surveillance Centre. SNOMED=Systematised Nomenclature of Medicine. *These items contain a proposed severity scale in RECAP-V0 and were captured in the electronic health records templates as a drop-down list of alternatives with appropriate per-item SNOMED coding; this ordering was carried through to the analysis. †Extracted from general practitioners electronic health records systems. ‡Collected from the caller as part of the NHS-111 pathway.

Table 2: Predictor variables considered for inclusion in the model and whether they were included following missingness assessment

R packages used for the analysis were rms, mice, miceMNAR, xplorerr, tidyverse, ggplot2, pubh, r2mlm, dplyr, tidyr, plotly, mlr3, and data.table.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

Between Oct 1, 2020, and Oct 26, 2021, 4278 patients (2415 [56.5%] from northwest London and 1863 [43.5%] from RCGP RSC) were recruited by 170 practices (103 [61%] in northwest London and 67 [39%] in RSC); CCAS enrolled 2674 patients; southeast London's Doctaly Assist provided records for 4033 patients (1948 [48.3%] for model development [Doctaly-1 dataset] and 2085 [51.7%] for model validation [Doctaly-2 dataset]; figure 1A and B). Although mean age and sex were similar across the four cohorts (table 1), there was a higher proportion of people minority ethnic groups in the northwest London and two Doctaly cohorts. We were

interested in assessing whether belonging to a minority ethnic group predicted higher risk of severity, as reported in previous publications.²² Initially, we had 16 ethnicity groups but, for statistical and policy making purposes, we agreed that grouping ethnicity categories into white and minority ethnic would produce more meaningful results.

The RCGP RSC data and the Doctaly-1 datasets were used for model development (figure 1B). The models were subsequently validated with the northwest London and Doctaly-2 data. Table 2 outlines all the model predictor variables considered and whether they were included in the model following assessment of patterns of missing data. Observations, such as SpO₂, were mostly missing in the northwest London, RSC, and CCAS data; however, SpO₂ was available for 1364 (70.0%) of 1948 patients who used Doctaly. Data on comorbidities were not available for the Doctaly Assist dataset because there was no linkage with GP health record data in southeast London.

Hospitalisation rates due to COVID-19 were similar in all the datasets except Doctaly-2: 83 (4.4%) of 1863 participants in the RSC, 92 (3.8%) of 2415 in

	Coefficient	p value
Intercept	-6.32	<0.0001
Sex		
Male	1 (ref)	..
Female	0.56	0.018
Age (years)	0.04	<0.0001
Hypertension history	0.56	0.04
Breathlessness		
Breathlessness (cannot complete sentences at rest)	1.69	<0.0001
Breathlessness on mild exertion	0.61	0.025
Breathlessness on moderate exertion	0.22	0.57*
Fever		
Temperature (rigors)	0.10	0.98*
Temperature (feeling feverish)	0.75	0.002

Absence of hypertension, breathlessness, and fever are the base coefficients in the logistic regression, set to zero and not shown. *For fever and breathlessness severity all levels were included if one level was significant. RECAP-GP=Remote COVID-19 Assessment in Primary Care-General Practice score.

Table 3: The RECAP-GP model

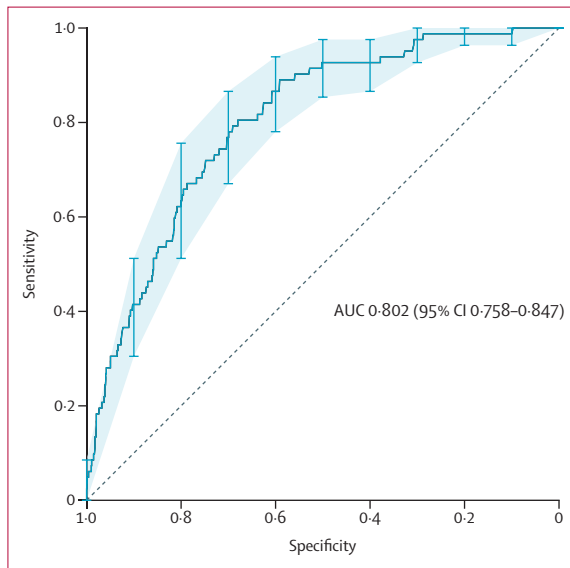


Figure 2: Receiver operating characteristic curve of the RECAP-GP model
 Bootstrapping for internal validation along with model diagnostic measures obtained as part of model calibration and performance assessment was done. Error bars and shaded areas are 95% CIs. AUC=area under the curve. RECAP-GP=Remote COVID-19 Assessment in Primary Care-general practitioner score.

northwest London, 82 (3.1%) of 2674 in CCAS, and 65 (3.3%) of 1948 in Doctaly-1, but only 19 (0.9%) of 2085 in Doctaly-2. All continuous data were found to be sufficiently normally distributed by visual inspection and the pattern of missingness for each variable was random. The pattern of data missingness was assumed to be missing at random because no clear pattern was observed in any of the variables considered for the model. The normal distribution of the variables allowed us to use multiple imputation chain equations for missing data imputation.

	Green risk group	Amber risk group	Red risk group
Risk group assigned by model in northwest London general practitioner data			
Actual hospitalisations	18 (1%) of 1453	50 (6%) of 797	22 (14%) of 158
Sensitivity (95% CI)	61.9% (59.9-63.9)	NA	24.4% (16.0-34.6)
Specificity (95% CI)	80% (70.2-87.7)	NA	94.1% (93.1-95.2)
Positive predictive value of red group designation (95% CI)	NA	NA	13.9% (9.8-19.4)
Negative predictive value of green group designation (95% CI)	98.8% (98.1-99.2)	NA	NA
Risk group assigned by model in CCAS data			
Actual hospitalisations	25 (2%) of 1512	45 (5%) of 958	12 (6%) of 204
Sensitivity (95% CI)	57.4% (55.4-59.3)	NA	14.6% (7.8-24.2)
Specificity (95% CI)	69.5% (54.4-79.2)	NA	93.0% (92.0-94.0)
Positive predictive value of red group designation (95% CI)	NA	NA	5.9% (3.5-9.7)
Negative predictive value of green group designation (95% CI)	98.3% (97.7-98.8)	NA	NA

Positive predictive value was calculated as the number of hospitalisations in red group divided by the number of patients in red group. Negative predictive value was calculated as the number of patients non-admitted in green group divided by the number of patients in green group. CCAS=COVID-19 Clinical Assessment Service. RECAP-GP=Remote COVID-19 Assessment in Primary Care-General Practice score. NA=not applicable.

Table 4: Validation of the RECAP-GP model

The RECAP-GP model was built with the RCGP RSC data. The predictor variables used in the final model were sex, age, history of hypertension, degree of breathlessness, and temperature symptoms (table 3). Fatigue, confusion, ethnicity, body-mass index (BMI), diabetes, coronary heart disease, and chronic kidney disease were excluded after backward elimination ($p>0.05$). Age and BMI were included as complex splines fitted through penalised thin plate regression. There was no non-linearity in the fitted splines because dimensionality was optimal at 1 degree of freedom, indicating a linear slope of fit. The model showed good performance to distinguish between risk levels (AUC 0.80 [95% CI 0.76-0.85]; figure 2). There was no substantial difference in the performance of the model when stratified for sex (0.81 [0.75-0.86] for men and 0.74 [0.68-0.81] for women) and age (0.70 [0.63-0.78] for patients ≥ 65 years and 0.73 [0.67-0.79] for those <65 years). The number of patients in the 65 years old or older group was lower (354 [19%] of 1863 patients) than the number of patients younger than 65 years (1509 [81%] patients), which was in line with the English population in 2019 (12.4 million [18.5%] of 67.1 million people).

	Coefficient	p value
Intercept	25.00	<0.0001
Age (years)	0.04	0.0002
Breathlessness		
Breathlessness (feeling uncomfortable to breathe)	0.92	0.030
Breathlessness when walking* around the room	-0.43	0.26
Fatigue		
Fatigue (difficult to wake up)	1.50	0.068
Fatigue (too tired to do usual activities)	1.23	0.0007
Oxygen saturation at rest (0–100%)	-0.33	<0.0001

Scales of severity are included if one element is significant and only the most severe level is used in the model. *For fatigue and breathlessness severity all levels were included if one level was significant. Absence of hypertension, breathlessness and fatigue are the base coefficients in the logistic regression, set to zero and not shown. RECAP-O2=Remote COVID-19 Assessment in Primary Care-oxygen saturation score.

Table 5: The RECAP-O2 model

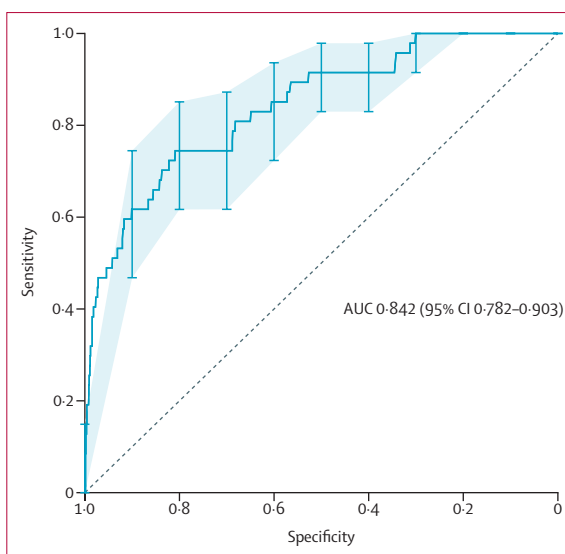


Figure 3: Receiver operating characteristic curve of the RECAP-O2 model
Bootstrapping for internal validation along with model diagnostic measures obtained as part of model calibration and performance assessment was done. Error bars and shaded areas are 95% CIs. AUC=area under the curve. RECAP-O2=Remote COVID-19 Assessment in Primary Care-oxygen saturation score.

The cutoff points for the green, amber, and red risk groups were chosen by the research team clinicians (BCD, SdL, ErM, EmM, ElM, TG, and AE-G) before validation, using the specificities and sensitivities obtained from the ROC (figure 2). We opted for maximising model sensitivity (90%) for the low-to-moderate risk threshold to ensure all patients who needed monitoring were in the amber group: the associated specificity was 40%, the logit transformed threshold 0.027, and the interval likelihood ratio (LR) 0.16. We maximised specificity (90%) at the moderate-to-high risk threshold to limit the number of unnecessary hospital admissions from the amber group, the associated sensitivity was 40%, the logit transformed threshold 0.098, and the interval LR 6.0.

For external validation, the prediction model was run using the northwest London data and the CCAS data separately. Following a data completeness assessment, seven (0.3%) of 2415 participants from northwest London cohort were removed, resulting in a sample size of 2408 with 90 hospitalisations. However, the data were similar to the GP data, and only 82 (3.1%) of 2674 patients were admitted to hospital, which was lower than initially expected admission rates and would have limited power for both model building and validation. Therefore, we used the CCAS data as a validation set only for the RECAP-GP model. We initially planned to use 1317 patients in the development set and 1400 patients in the for-validation set with 10% admitted and a 0.05 margin of error.¹³ The selected cutoff points were used to assign risk categories to patients (table 4), along with the observed model sensitivity, specificity, negative predictive value, and positive predictive value. Because true negative and true positive cannot be defined for the amber group, only the number of hospitalisations in this group is reported. In the northwest London GP data, the

probability of being categorised as a patient at low risk (green) and not needing admission (ie, negative predictive value) was high (1435 [99%] of 1453; 95% CI 98–99), and the probability of being in the high risk (red) group and being admitted (ie, positive predictive value) was low (22 [14%] of 158; 10–19). In the CCAS data, the negative predictive value was (1487 [98%] of 1512; 95% CI 98–99), equivalent to the GP data, but the positive predictive value was lower (12 [6%] of 204; 4–10).

Predictor variables used in the final RECAP-O2 model were age, degree of breathlessness, fatigue, and SpO₂ at rest (table 5). Sex, ethnicity, temperature, acute cognitive decline, days since onset of symptoms, respiratory rate, and trajectory of breathlessness were excluded after backwards elimination ($p>0.05$). SpO₂ after activity was found to be colinear with SpO₂ at rest in the model and was thus excluded from the final model. To explore the potential non-linearity of continuous parameters (age and oxygen saturation at rest) these were modelled as complex splines fitted through penalised thin plate regression. There was no non-linearity found in age, and slight non-linearity found in oxygen saturation at rest, estimated as 1.7 degrees of freedom. The degree of non-linearity was assessed by varying the number of the basis dimension used for spline fitting, and inspecting its influence on the effective number of dimensions in the final fit, with no significant change found. The effect of adding oxygen saturation at rest as a non-linear term in the model was investigated and no significant improvement in performance was found, with a negligible improvement in AIC of 0.15 and no change in McFadden's pseudo R².

	Green risk group	Amber risk group	Red risk group
Actual hospitalisations	7 (1%) of 1183	9 (2%) of 457	3 (9%) of 34
Sensitivity (95% CI)	71.1% (68.8–73.2)	NA	15.8% (3.4–39.6)
Specificity (95% CI)	63.1% (38.4–83.7)	NA	98.1% (97.3–98.7)
Positive predictive value of red group designation (95% CI)	NA	NA	8.8% (3.1–22.4)
Negative predictive value of green group designation (95% CI)	99.4% (98.9–99.7)	NA	NA

Positive predictive value was calculated as the number of hospitalisations in red group divided by the number of patients in red group. Negative predictive value was calculated as the number of patients non-admitted in green group divided by the number of patients in green group. RECAP-O2=Remote COVID-19 Assessment in Primary Care-oxygen saturation score. NA=not applicable.

Table 6: Validation of the RECAP-O2 model Risk group assigned by model in Doctaly-2 validation data

The model was internally validated and calibrated through bootstrapping using 500 repeats. The AUC was 0.84 (95% CI 0.78–0.90; figure 3), which suggests good model performance. Of note, the bottom left of the ROC curve has a steeper slope than the RECAP-GP model, indicating better discrimination in patients with more severe disease. There was no significant difference in the performance of the model when stratified for sex (AUC 0.86 [95% CI 0.78–0.95] for men and 0.83 [0.74–0.91] for women) and age (0.86 [0.75–0.98] for patients ≥ 65 years and 0.83 [0.77–0.90] for those < 65 years),

External validation the RECAP-O2 model was run using the Doctaly-2 data. Following data completeness assessment, 411 (19.7%) of 2085 participants were removed from the Doctaly-2 sample, resulting in a sample size of 1674. The cutoff points were chosen by the clinical team based on the ROC curve (figure 3) and were used to assign risk categories to patients. The green–amber risk cutpoint at 90% sensitivity had a specificity of 52%, a logit transformed threshold of 0.013 and an interval LR of 0.19. For the amber–red risk cutpoint, 95% specificity had a sensitivity of 48%, a logit transformed threshold of 0.096, and an interval LR of 10.4. Validation results are shown in table 6. Although the negative predictive value of patients in the green risk group was slightly higher than in RECAP-GP (1176 [99%] of 1183 participants; 95% CI 99–100), positive predictive value of patients in the red risk group was lower (three [9%] of 34 participants; 95% CI 3–22), which means lower probability of being admitted when categorised as high risk. Performance stratification by age and sex was also assessed in the external validation dataset. However, this was limited by the low number of admissions because 18 (1%) of 2085 participants in Doctaly-2 were hospitalised and there was only one hospitalisation in the subgroup of patients aged 65 years old or older. The results from Doctaly-2 were not representative of the model (appendix p 19).

Discussion

Assessment of severity of COVID-19 in the community is crucial to pandemic management worldwide. Our study

provides a derivation and real-world validation of two risk scores specifically designed for patients with COVID-19 in the community. The RECAP-GP model includes degree of breathlessness, temperature symptoms, history of hypertension, sex, and age as hospital admission predictors, and can be used when a pulse oximeter is not available. This model provides a good prediction for non-admission in the lowest risk group. The model performs less well at differentiating amber from red risk groups. When validated in the CCAS data the positive predictive value is slightly lower but performance on the green and amber risk cutoff point is within the northwest London CIs. The RECAP-O2 model included degree of breathlessness, fatigue, SpO₂ at rest, and age as predictors. This model can be used if pulse oximeters are available, including in patients at moderate risk who are being monitored. Although the improvement of the slope on the ROC curves indicates an interval LR increasing from 6 to 10 for the amber to red risk cutpoint, the model performs less well on validation. However, its specificity is good (98% [95% CI 97–99]), which supports its use to assess patients at moderate risk.

The RECAP models establish an evidence base for the assessment of patients with COVID-19 in the community.^{5,23} A recent systematic review identified that monitoring of SpO₂ in the community (both at rest and after exercise) was useful to identify risk of patient deterioration, with an SpO₂ of 92–94% considered the lower threshold for treatment escalation.²³ In our study, an SpO₂ of 92% placed the patient at the threshold for the high risk group. There was collinearity of SpO₂ at rest and after exertion. Thus, only SpO₂ at rest is used in the RECAP-O2 model; however, degree of breathlessness, when shortness of breath after movement is assessed, is included as a predictor factor. In the UK a national strategy for using home pulse oximetry, COVID Oximetry @home, was established in 2020 to identify silent hypoxia (hypoxia without breathlessness). UK guidance recommends provision of pulse oximeters to monitor patients with symptomatic COVID-19 and individuals older than 65 years or with specific long term conditions putting them at risk of severe COVID-19 (eg, severe liver or kidney disease).^{19,24} The use of RECAP in the assessment of patients with COVID-19 is a valuable addition to this strategy. Our results are aligned with the UK guidance since age and SpO₂ significantly predict deterioration in our models; additionally, RECAP considers symptom severity, so it can better support clinicians' judgement on who needs monitoring, particularly for younger patients without comorbidities. Moreover, RECAP-O2 is better at identifying need for treatment escalation compared with SpO₂ alone, which is only one factor in the final model.

The use of multiple datasets to develop the models is a strength of the study. Moreover, the age distributions, admission rates, ethnicity, and comorbidities are in line with UK population expectations.^{25,26} The cohorts from southeast and northwest London contain larger

populations of Black British and South Asian ethnicity, whereas the RSC network included a smaller proportion of participants from minority ethnic groups.²⁵ This supports the external generalisability of our findings. The thresholds used in the two RECAP scores can be adjusted to better suit local circumstances. Although the RECAP-V0 template, developed by the Delphi study, contained ten questions,¹² the validated models contained four (RECAP-O2) and five (RECAP-GP) items only, significantly improving their fitness for use in the clinical setting.²⁷ The choice of logistic regression modelling means that most electronic health record systems using SNOMED codes will be able to recreate the RECAP electronic templates and integrate the score calculator into the system.²⁸ Moreover, data on hospitalisation was extracted from the NHS databases, which contain information on admission to tertiary NHS care facilities in England and are considered a reliable data source. Even if a patient initially assessed in northwest London was hospitalised in a different region or NHS area, their information would have been captured, and the data linked, in the databases used. The full models are provided as downloadable code on Github.

The study has limitations. The RECAP models are founded on assessment of need for hospital admission in an observational dataset. In such a design it is impossible to completely eliminate incorporation bias. We mitigated this by counting admissions as at least one night in hospital, rather than only a review in the emergency department. Admission overnight will be based on investigation in hospital or a need for hospital-based therapy, thus reducing the role of RECAP score elements in the clinical decision that is the outcome variable. Health monitoring devices, such as pulse oximeters, are rarely available in the community. Therefore, these observations were largely missing in the GP and CCAS datasets and could not be included in the RECAP-GP model. By contrast, the Doctaly Assist datasets enabled us to assess the predictive value of resting and after exercise SpO₂, heart rate, respiratory rate, and temperature, determining, for the first time, the diagnostic value of SpO₂ monitoring in the community. The poor availability of electronic health record data linkage for the Doctaly data prevented the inclusion of comorbidities in the RECAP-O2 model, which might have contributed to the less good calibration of the model at higher risk scores.

Moreover, although most of the recruitment was done from October, 2020, to May, 2021, the Doctaly Assist validation dataset was collected from May to October, 2021, at which point the UK's COVID-19 vaccination programme covered more than 70% of the adult population (>18 years).²⁹ The potential effect of vaccination in hospital admission, along with the possible empirical treatment of COVID-19 in the community,^{30,31} and the lower mean age of the Doctaly-2 cohort compared with other cohorts might explain the

lower number of hospitalisations in this cohort and the difference in the RECAP-O2 model validation (including the lower performance when assessing performance by age and sex subgroups in the external validation dataset; appendix p 20) because lower admission rates require even better model discrimination to achieve good positive predictive value, and require a larger sample size than we planned.

At the time of data collection, vaccination status was not consistently available in the electronic health record and policy was changing rapidly during the study. Our models are based on the associations between symptoms severity, and patient's demographics and comorbidities, and need for hospital admission (primary outcome). As recent publications report,³² vaccination status might affect the severity of symptoms, but the relationship between symptoms severity and likelihood of hospitalisation is not expected to change. Similarly, severity of symptoms in infections caused by different SARS-CoV-2 variants might also differ, but this will probably not change the relationship between severity of symptoms and admission probability. We believe that our models are relevant to assess the likelihood of admission based on the patient's symptoms, regardless of the variant identified and vaccination status. However, all models should be subject to ongoing surveillance and calibration to ensure factors considered remain relevant, especially with rapid changes in variants and vaccines. Moreover, vaccination rate and treatment availability in different settings might invoke a change in the population baseline risk of hospitalisation, which might affect the absolute risk (while probably preserving the prediction power of model factors). This might justify recalibration and re-estimation of model intercept in populations with potentially different baseline risk in future studies. Depending on data availability, recalibration could involve the collection of data on predictors and predicted factors in a new population and additional information on vaccination status, PCR result, including type of variant, and COVID-19 treatment received in the community, to assess their predictive power. Despite the need for continuing calibration, the RECAP-GP model showed a good performance for all risk levels, which supports its use by practitioners to decide the need for monitoring patients with signs and symptoms of COVID-19. We suggest that the RECAP-GP score be used remotely in the initial assessment, without need for patient observations. Following this, if the patient is considered moderate to high risk, they could be provided with a home pulse oximeter for calculation of the RECAP-O2 score to detect deterioration. Given the low positive predictive value of the red categorisation on the amber to red cutpoint, the RECAP models will tend to overalert or overestimate (particularly for risk for higher risk groups). Care escalation should be considered with reference to national or local pathways, ability to monitor the patient

in the community, hospital capacity, and shared decision making with the patient.

Much has evolved since the first wave of the COVID-19 pandemic when we conceptualised this study. Mass vaccination has dramatically changed prognosis and oximetry is now much more widely used in the community than it was in early 2020, with at-home services available in many settings. Yet, new variants are triggering new pandemic waves across the globe, putting services under great strain. We believe that these two scores are likely to be valuable resources to support clinical judgement, reduce uncertainty, and improve safety in triage and monitoring of patients with suspected COVID-19 in health systems worldwide.

Contributors

AE-G coordinated recruitment across sites, helped with template development and elaboration of datasets in northwest London Whole Systems Integrated Care/ Imperial clinical, analytics, research, and evaluation (iCARE) and ORCHID, and drafted the manuscript. DP and FF designed and did the analysis and interpreted the results. CR supported the ethics submission and amendments, communications with practices during the recruitment period, Data Protection Impact Assessment, Control of Patient Information, and Confidentiality Advisory Group applications. EIM and EmM developed the code and template, and analysed and interpreted the results. BG conceptualised and designed the study, managed the data in iCare, analysed and interpreted the results. ALN developed the protocol, obtained funding, and interpreted the results. CO managed the data in ORCHID and analysed interpreted the results. LH obtained funding, wrote the protocol, and managed data collection. RC and JH designed the South East London COVID management pathways, obtained funding, collected the data, and interpreted the analyses. MB, CW, and BB coordinated the patient enrolment at the COVID-19 Clinical Assessment Service. BCD, ErM, SdL, and TG developed the protocol and obtained funding. BCD supervised the study. All authors had access to the summary data and contributed to interpretation of results and drafting of the manuscript. AE-G, EIM, EmM, and DP verified the data in the iCare and ORCHID secure environments and produced the data summaries, as data governance restricts full data access to a restricted number of researchers.

Declaration of interests

SdL is Director of the Royal College of General Practitioners Research and Surveillance Centre, the English primary care sentinel system. This work funds part of his academic position, and the network and its Trusted Research Environment were part of this study, funded using the university standard cost template. SdL reports grants through his University from AstraZeneca, Eli Lilly, GSK, Sanofi, Seqirus, and Takeda, none have a direct link to this study and has been a member of Advisory Boards for AstraZeneca, Sanofi, and Seqirus, none have a direct link to this study. All other authors declare no competing of interests.

Data sharing

Due to data governance limitations, the deidentified patient data used to develop and validate the models cannot be shared. The model code is available in on Github.

Acknowledgments

Doctaly Assist (Phil Tyler and Prad Velayuthan, Doctaly) provided the data extraction and Systematized Nomenclature of Medicine (SNOMED) coding of their chat-bot data. Mark Ashworth and Ibidun Fakoya (King's College London, London, UK), evaluated the initial set-up of the COVID monitoring programme in Southeast London. Eamon O'Doherty (Northwest London Clinical Commissioning Group, London, UK) extracted the relevant data tables from northwest London Whole Systems Integrated Care to enable the data analysis in iCare. Sneha Anand (Oxford University, Oxford, UK) managed the governance, data transfer, and linkage in ORCHID. Matt Widdows (South Central Ambulance Service NHS Trust) coordinated the

development of the RECAP template with Adastra and extracted the relevant CCAS data for analysis. Merlin Dunlop (Ardens, Oxford, UK) assisted in the SNOMED coding of the RECAP-V0 templates and their distribution. Ashnee Dhondee (northwest London Clinical Research Network, London, UK) helped recruit practices to the study. Patients and practices of the Royal College of General Practitioners Research and Surveillance Centre and other networks shared data for this research. All those acknowledged have seen a copy of the final draft paper. The study was funded by the Community Jameel and the Imperial College President's Excellence Fund, the Economic and Social Research Council, the UK Research and Innovation, and Health Data Research UK. The authors gratefully acknowledge infrastructure support from the National Institute for Health and Care Research (NIHR) Imperial Patient Safety Translational Research Centre, the NIHR Imperial Biomedical Research Centre and the NIHR Oxford Biomedical Research Centre. This research was in part enabled by the Imperial Clinical Analytics Research and Evaluation (iCARE) environment and Whole System Integrated Care and used the iCARE and WSIC team and data resources. The project was supported by the NIHR CRN Urgent Public Health Study.

References

- Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. *BMJ* 2020; **368**: m1182.
- Nunan D, Brassey J, Mahtani K, et al. COVID-19 signs and symptoms tracker. Oxford. 2020. <https://www.cebm.net/COVID-19/COVID-19-signs-and-symptoms-tracker/> (accessed Dec 18, 2021).
- Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* 2020; **323**: 1239–42.
- 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; **369**: m1328.
- NHS England. Pulse oximetry to detect early deterioration of patients with COVID-19 in primary and community care settings. 2020. <https://www.england.nhs.uk/coronavirus/publication/pulse-oximetry-to-detect-early-deterioration-of-patients-with-covid-19-in-primary-and-community-care-settings/> (accessed Jan 12, 2021).
- Knight SR, Ho A, Pius R, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO clinical characterisation protocol: development and validation of the 4C Mortality Score. *BMJ* 2020; **370**: m3339.
- Institute of Social and Preventive Medicine. Living evidence on COVID-19. 2022. <https://ispmbern.github.io/covid-19/living-review/index.html> (accessed Jun 28, 2022).
- 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.
- Carr E, Bendayan R, Bean D, et al. Evaluation and improvement of the national early warning score (NEWS2) for COVID-19: a multi-hospital study. *BMC Med* 2021; **19**: 23.
- Hodgson LE, Congleton J, Venn R, Furni LG, Roderick PJ. NEWS 2 – too little evidence to implement? *Clin Med* 2018; **18**: 371–73.
- Tankel JW, Ratcliffe D, Smith M, et al. Consequences of the emergency response to COVID-19: a whole health care system review in a single city in the United Kingdom. *BMC Emerg Med* 2021; **21**: 55.
- Greenhalgh T, Thompson P, Weiringa S, et al. What items should be included in an early warning score for remote assessment of suspected COVID-19? qualitative and Delphi study. *BMJ Open* 2020; **10**: e042626.
- Espinosa-Gonzalez AB, Neves AL, Fiorentino F, et al. Predicting risk of hospital admission in patients with suspected COVID-19 in a community setting: protocol for development and validation of a multivariate risk prediction tool. *JMIR Res Protoc* 2021; **10**: e29072.
- Fiorentino F, Prociuk D, Espinosa Gonzalez AB, et al. An early warning risk prediction tool (RECAP-V1) for patients diagnosed with COVID-19: protocol for a statistical analysis plan. *JMIR Res Protoc* 2021; **10**: e30083.

For the model code see <https://github.com/DenysProciuk/RECAP>

- 15 de Lusignan S, Dorward J, Correa A, et al. Risk factors for SARS-CoV-2 among patients in the Oxford Royal College of general practitioners research and surveillance centre primary care network: a cross-sectional study. *Lancet Infect Dis* 2020; **20**: 1034–42.
- 16 de Lusignan S, Lopez Bernal J, Zambon M, et al. Emergence of a novel coronavirus (COVID-19): protocol for extending surveillance used by the Royal College of General Practitioners Research and Surveillance Centre and Public Health England. *JMIR Public Health Surveill* 2020; **6**: e18606.
- 17 Doctaly. Doctaly Assist. 2021. <https://doctalyassist.com/lewisham/covid> (accessed 11 December 2021).
- 18 WhatsApp. WhatsApp. 2020. <https://whatsapp.com> (accessed March 30, 2022).
- 19 NHS. COVID Oximetry @home. 2021. <https://www.england.nhs.uk/nhs-at-home/covid-oximetry-at-home/> (accessed Dec 11, 2021).
- 20 Clarke J, Flott K, Fernandez Crespo R, et al. Assessing the safety of home oximetry for COVID-19: a multisite retrospective observational study. *BMJ Open* 2021; **11**: e049235.
- 21 Ramos-Casals M, Brito-Zerón P, Mariette X. Systemic and organ-specific immune-related manifestations of COVID-19. *Nat Rev Rheumatol* 2021; **17**: 315–32.
- 22 Bray I, Gibson A, White J. Coronavirus disease 2019 mortality: a multivariate ecological analysis in relation to ethnicity, population density, obesity, deprivation and pollution. *Public Health* 2020; **185**: 261–63.
- 23 Alboksmaty A, Beaney T, Elkin S, et al. Effectiveness and safety of pulse oximetry in remote patient monitoring of patients with COVID-19: a systematic review. *Lancet Digit Health* 2022; **4**: e279–89.
- 24 NHS. Who is at high risk from coronavirus (COVID-19). 2022. <https://www.nhs.uk/conditions/coronavirus-covid-19/people-at-higher-risk/who-is-at-high-risk-from-coronavirus/> (accessed June 29, 2022).
- 25 Office for National Statistics. Population estimates for the UK, England and Wales, Scotland and Northern Ireland: mid-2019. 2020. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/bulletins/annualmidyearpopulationestimates/mid2019estimates> (accessed Jan 9 2022).
- 26 NHS Digital. Quality and Outcomes Framework, 2020–21. 2021. <https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/2020-21/main-findings> (accessed April 2, 2022).
- 27 Robert A. Jenders, Decision rules and expressions. In: Greenes RA, ed. *Clinical decision support*. Cambridge, MA: Academic Press, 2014: 417–34.
- 28 Correa A, Hinton W, McGovern A, et al. Royal College of General Practitioners Research and Surveillance Centre (RCGP RSC) sentinel network: a cohort profile. *BMJ Open* 2016; **6**: e011092.
- 29 UK Government. Coronavirus (COVID-19) in the UK. 2021. <https://coronavirus.data.gov.uk/details/vaccinations> (accessed Dec 12, 2021).
- 30 NHS. Interim Clinical Commissioning Policy: antivirals or neutralising monoclonal antibodies for non-hospitalised patients with COVID-19 (version 5). 2022. <https://www.england.nhs.uk/coronavirus/publication/interim-clinical-commissioning-policy-neutralising-monoclonal-antibodies-or-antivirals-for-non-hospitalised-patients-with-covid-19/> (accessed July 11, 2022).
- 31 Yu L-M, Bafadhel M, Dorward J, et al. Inhaled budesonide for COVID-19 in people at high risk of complications in the community in the UK (PRINCIPLE): a randomised, controlled, open-label, adaptive platform trial. *Lancet* 2021; **398**: 843–55.
- 32 Antonelli M, Penfold RS, Merino J, et al. Risk factors and disease profile of post-vaccination SARS-CoV-2 infection in UK users of the COVID Symptom Study app: a prospective, community-based, nested, case-control study. *Lancet Infect Dis* 2022; **22**: 43–55.