BMJ Open

Development and validation of an early warning score to identify COVID-19 in the emergency department based on routine laboratory tests: a multicentre case–control study

Arjen-Kars Boer,[1] Ruben Deneer ,[1,2] Maaike Maas,[3] Heidi S M Ammerlaan,[4] Roland H H van Balkom,[5] Wendy A H M Thijssen,[6] Sophie Bennenbroek,[6] Mathie Leers ,[7] Remy J H Martens,[7] Madelon M Buijs,[8] Jos J Kerremans,[9] Muriël Messchaert,[10] Jeroen J van Suijlen,[10] Natal A W van Riel,[2,11] Volkher Scharnhorst[1,2]

To cite: Boer A-K, Deneer R, Maas M, *et al*. Development and validation of an early warning score to identify COVID-19 in the emergency department based on routine laboratory tests: a multicentre case–control study. *BMJ Open* 2022;**12**:e059111. doi:10.1136/bmjopen-2021-059111

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-059111).

A-KB and RD contributed equally.

Received 10 November 2021
Accepted 10 June 2022

For numbered affiliations see end of article.

Correspondence to
Ruben Deneer;
ruben.deneer@catharinaziekenhuis.nl

ABSTRACT

Objectives Identifying patients with a possible SARS-CoV-2 infection in the emergency department (ED) is challenging. Symptoms differ, incidence rates vary and test capacity may be limited. As PCR-testing all ED patients is neither feasible nor effective in most centres, a rapid, objective, low-cost early warning score to triage ED patients for a possible infection is developed.

Design Case–control study.

Setting Secondary and tertiary hospitals in the Netherlands.

Participants The study included patients presenting to the ED with venous blood sampling from July 2019 to July 2020 (n=10 417, 279 SARS-CoV-2-positive). The temporal validation cohort covered the period from July 2020 to October 2021 (n=14 080, 1093 SARS-CoV-2-positive). The external validation cohort consisted of patients presenting to the ED of three hospitals in the Netherlands (n=12 061, 652 SARS-CoV-2-positive).

Primary outcome measures The primary outcome was one or more positive SARS-CoV-2 PCR test results within 1 day prior to or 1 week after ED presentation.

Results The resulting 'CoLab-score' consists of 10 routine laboratory measurements and age. The score showed good discriminative ability (AUC: 0.930, 95% CI 0.909 to 0.945). The lowest CoLab-score had high sensitivity for COVID-19 (0.984, 95% CI 0.970 to 0.991; specificity: 0.411, 95% CI 0.285 to 0.520). Conversely, the highest score had high specificity (0.978, 95% CI 0.973 to 0.983; sensitivity: 0.608, 95% CI 0.522 to 0.685). The results were confirmed in temporal and external validation.

Conclusions The CoLab-score is based on routine laboratory measurements and is available within 1 hour after presentation. Depending on the prevalence, COVID-19 may be safely ruled out in over one-third of ED presentations. Highly suspect cases can be identified regardless of presenting symptoms. The CoLab-score is continuous, in contrast to the binary outcome of lateral flow testing, and can guide PCR testing and triage ED patients.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ A comprehensive panel of 28 laboratory tests were measured for 10 417 emergency department (ED) presentations and combined with SARS-CoV-2 PCR test results.
⇒ Using adaptive lasso regression analysis, the panel of 28 laboratory tests was reduced to a single score consisting of a subset of 10 routine ED laboratory tests and age.
⇒ The score was temporally validated from July 2020 to October 2021, in the presence of vaccine roll-out and emergence of new SARS-CoV-2 variants.
⇒ The score was externally validated in three other centres in the Netherlands.
⇒ Missingness in the panel of laboratory tests varied between the external centres, limiting generalisability of the score to the ED population for which the complete panel of laboratory tests was available.
⇒ The score was not directly compared with lateral flow testing.

INTRODUCTION

COVID-19, caused by SARS-CoV-2, evolved into a global pandemic in 2020.[1] For emergency department (ED) physicians, identifying presenting patients with a possible COVID-19 infection remains challenging since symptoms like fever, shortness of breath or coughing overlap with other illnesses.[2 3] It is crucial, however, to identify a possible COVID-19 infection as early as possible. Early identification prevents further spreading and protects hospital staff by isolating a suspected patient, pending the results of a SARS-CoV-2 RNA PCR test and/or chest CT. Conversely, when PCR testing or isolation treatment

capacity is limited, ruling out COVID-19 as soon as possible can save valuable resources.

In the era of electronic health records and clinical prediction models, developing an early warning score that can assist ED physicians in identifying patients presenting to the ED with COVID-19 is of great value. Moreover, if only routine ED test results are required as input, the score can be easily adopted by EDs worldwide, potentially reduce diagnostic costs and accelerate patient triage.

Many COVID-19 prediction models have already been developed; the living systematic review by Wynants *et al*[4] provides an extensive overview and critical appraisal. Unfortunately, only few models have found their way into routine care at the ED.[5 6] Early models were based on relatively small sample sizes, hampered by selection bias or were overfitted by selecting too many features.[4–6] Aside from the methodological shortcomings of early models, most models are not developed as an early warning score for all ED patients. First, they require features from tests that are not routinely performed or logged for all ED patients (eg, the COVID-19 Reporting and Data System-score from a CT scan[7] or non-laboratory-based clinical variables in the Pandemic Respiratory Infection Emergency System Triage Early Warning Score)[8] and are therefore not straightforward to implement or scale to a large ED patient population. Second, the population on which models are commonly based are PCR-tested patients, that is, a preselection of a possible COVID-19 infection has already been done by physicians.

Only two studies were identified that focus on patients presenting to the ED, include unsuspected (and prepandemic) patients as controls and rely solely on routine (laboratory) tests.[9 10]

In this study we report the development and validation of an early warning score that, based on routine ED laboratory tests, estimates the risk of a possible COVID-19 infection in patients who undergo routine laboratory testing at presentation. The score can assist ED physicians in triaging patients and prevent further transmission of COVID-19 by quickly identifying possibly infected patients or ruling out a possible infection when resources are scarce.

METHODS

Study design

This is a retrospective case–control study where routine laboratory test results, combined with age and gender, from all patients presenting to the ED of the Catharina Hospital Eindhoven from July 2019 to July 2020 were combined with SARS-CoV-2 PCR test results in a development data set. A model that could predict the presence of a COVID-19 infection was fit to this data set. The performance of the model was assessed by (1) internal validation, (2) temporal validation and (3) external validation by using data from the ED of three other centres.

Patient and public involvement

Patients were not involved in the design, conduct or reporting of this study.

Development data set

All ED presentations at the Catharina Hospital Eindhoven from July 2019 to July 2020 were included in the development data set, provided that routine laboratory testing had been requested by the attending ED physician. The rationale for this inclusion period is to limit the effect of seasonal variation in the ED patient population by including the summer, fall and winter seasons of 2019 (control patients) and the winter, spring and summer seasons of 2020 (case and control patients). The routine laboratory panel at the ED consists of 28 laboratory tests. In some cases not all tests in the routine panel were requested or one or more quantitative results were not available due to analytical interference (haemolysis, lipaemia or icterus). The routine ED laboratory panel is requested for (adult) patients presenting with abdominal pain, chest pain, shortness of breath, syncope, sepsis or other non-specific complaints, or for patients (including non-adult patients) presenting with specific complaints where a suspected diagnosis has to be ruled in or ruled out. Presentations with one or more missing values in any of the 28 laboratory tests in the routine ED panel were excluded. Presentations with one or more extreme laboratory results, >10 times the SD from the median, were also excluded to minimise the effect on the estimation of regression coefficients. The median was chosen as a measure of central tendency due to its resistance for outliers. After the first case of COVID-19 in the Netherlands, all patients with symptoms of COVID-19 (either fever and/or respiratory symptoms) were subjected to nasopharyngeal PCR testing for SARS-CoV-2 RNA. PCR testing was performed by commercial tests that were approved by the Dutch National Institute of Public Health (RIVM). If a patient had a positive PCR result in the past, subsequent presentations were excluded as re-presentations might be clinically different from de novo presentations.

The ED laboratory panel results were matched to SARS-CoV-2 PCR results if the underlying nasopharyngeal swab had been taken ≤1 day prior or ≤1 week after initial blood withdrawal at the ED. If multiple PCR tests were performed in this window and at least one PCR test was positive, the presentation was labelled 'PCR-positive'. If all PCR test results in the time window were negative, the presentation was labelled as 'PCR-negative'. If no PCR tests were performed in the time window and the presentation occurred after the first case of COVID-19 in the Netherlands, the presentation was labelled as 'Untested'. All presentations before the first case were labelled as 'Pre-COVID-19'.

Laboratory tests

The routine laboratory panel consisted of haemocytometric and chemical analyses. The haemocytometric tests

were performed on Sysmex XN-10 instruments (Sysmex, Kobe, Japan) and consisted of haemoglobin, haematocrit, erythrocytes, mean corpuscular volume, mean cellular haemoglobin, mean cellular haemoglobin concentration, thrombocytes, leucocytes, neutrophils, eosinophils, basophils, lymphocytes and monocytes. The chemical analyses were performed on a Cobas 8000 Pro (Roche Dx, Basel, Switzerland) instrument and consisted of glucose, total bilirubin, aspartate aminotransferase (ASAT), alanine aminotransferase (ALAT), lactate dehydrogenase (LD), creatine kinase (CK), alkaline phosphatase (ALP), gamma-glutamyltransferase (gGT), blood urea nitrogen, creatinine, chronic kidney disease epidemiology collaboration (CKD-epi) estimated glomerular filtration rate (eGFR), potassium, sodium, chloride, albumin (bromocresol green) and C reactive protein (CRP). These results were combined with age and gender.

Modelling

All data were processed and analysed in R V.4.1.1.[11] Laboratory results, combined with age and gender, were used as covariates in the regression model. Cases were defined as ED presentations labelled as 'PCR-positive'; controls were all other presentations (ie, 'PCR-negative', 'Untested' or 'Pre-COVID-19'). To achieve predictive accuracy, limit overfitting and perform feature selection, penalised logistic regression with an adaptive lasso penalty was chosen.[12 13] To minimise missing data, all non-numeric results at the extremes of the measuring range were converted to numeric results by removing the '<' and '>' signs. For eGFR (CKD-epi) and CRP the raw precursor value was used instead of >90 mL/min/m^2 and <6 mg/L, respectively. Considering that laboratory results of bilirubin, ASAT, ALAT, LD, CK, ALP and gGT can have heavy (right) tailed distributions, which in turn impact model predictions, these variables were transformed logarithmically. More details regarding model fitting can be found in online supplemental material 1. Models were fitted using the glmnet package.[14]

CoLab-score

Since this is a retrospective case–control study, the sample prevalence may not reflect the true/current COVID-19 prevalence. To obtain well-calibrated probabilities, the intercept term in the model should be adjusted according to the current prevalence (details can be found in online supplemental material 1).[15] However, adjusting the intercept term is not straightforward to implement in clinical practice; therefore, the linear predictor of the model was categorised into a score and this score is hereafter referred to as the 'CoLab-score'. The categorisation is based on a number needed to test of 15 (ie, one is willing to PCR-test 15 patients to find one positive) and prevalence cut-points of 1%, 2%, 5%, 10% and 40% using the intercept adjustment formula by King and Zeng.[15] The intervals obtained through these breaks correspond to CoLab-scores 5 to 0, respectively. A score of 0 reflects low risk for COVID-19 and a score of 5 reflects high risk. More details regarding the rationale of the CoLab-score categorisation can be found in online supplemental material 1.

Internal validation

To assess model performance while taking overfitting into account, bootstrapping was performed. From the original data, 1000 bootstrap samples were generated. On each bootstrap sample, full model fitting procedure and CoLab-score conversion were performed. Optimism-adjusted performance measures of the CoLab-score were obtained by applying the 0.632 bootstrap rule to the in-sample and out-of-bag-sample performance.[16] Performance measures included the area under the ROC-curve (AUC), sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of each CoLab-score. The pROC package was used to calculate the performance measures.[17] Although the full inclusion period from July 2019 to July 2020 was used for model fitting, the performance was evaluated on the period starting from the first COVID-19 infection (24 February 2020) to July 2020. This was done to obtain performance measures that would reflect real-world performance.

Temporal validation

For temporal validation, the results from our centre were prospectively analysed from July 2020 to October 2021. During this period, the Netherlands was struck by a second wave of COVID-19 infections, starting in the fall of 2020 and subsiding in the summer of 2021. In this period there was also more widespread external PCR testing by municipal health services. The results of external conducted PCR tests were not available for our study. To overcome this limitation, the outcome in the temporal validation cohort was chosen as a composite of the hospital registration of a confirmed COVID-19 infection and/or at least one positive PCR test result. This period also covers both the emergence of new SARS-CoV-2 variants as well as vaccine roll-out. However, neither vaccination status nor genomic sequencing was available to determine whether a patient was vaccinated or which variant caused the infection. Therefore, data from the Dutch National Institute of Public Health (RIVM) were used to divide the temporal validation period into three phases: (1) from July 2020 until March 2021, no vaccination and no variants of concern identified; (2) from March 2021 until June 2021, partial vaccination and B.1.1.7 (Alpha) variant identified as dominant; and (3) from June 2021 until October 2021, widespread vaccination and B.1.617.2 (Delta) variant identified as dominant. See figure 1 in online supplemental material 2 for more details. The temporal validation consisted of assessment of the AUC, sensitivity, specificity, PPV and NPV of each CoLab-score threshold for the entire period, as well as for each phase separately to determine a possible effect of vaccination and new variants on performance (results in online supplemental material 2). Model calibration was assessed graphically using the rms package.[18]

External validation

For the external validation, several centres in the Netherlands were approached and assessed if the required panel of laboratory tests and SARS-CoV-2 PCR test results were available. Seven centres responded and three centres fulfilled the inclusion criteria: Gelre Hospitals (centre 1), Atalmedial Diagnostic Centers, location Alrijne Hospital Leiderdorp (centre 2) and Zuyderland Medical Center (centre 3). The haematological parameters were measured with Sysmex XN-10/XN-20 (centre 1), CELL-DYN Sapphire (Abbott Laboratories) (centre 2) and Sysmex XN-10 instruments (centre 3). The clinical chemistry parameters were measured with Architect c14100/c160000 (Abbott Laboratories) (centre 1), Architect ci4100 (Abbott Laboratories) (centre 2) and Cobas 8000 instruments (Roche Dx) (centre 3). The external validation was similar to the temporal validation and consisted of assessment of the AUC, sensitivity, specificity, PPV and NPV of each CoLab-score threshold. Calibration was assessed graphically analogous to the temporal validation data set.

RESULTS

Development data set

The study included 12 879 ED presentations of 10 327 patients from July 2019 to July 2020. After excluding cases with an incomplete laboratory panel, patient presentations that occurred after a positive PCR test in the past (re-presentations) and presentations with extreme values (>10 times SD) in any of the laboratory results, 10 417 presentations of 8610 patients remained (figure 1A).[19]

Descriptive statistics of ED presentations are shown in table 1, where the symbol '‡' indicates a clinically relevant difference from the pre-COVID-19 category (based on the total allowable error[20]). For the PCR positives (n=279), 91% (95% CI 88% to 94%) of the cases were tested positive in their first PCR. The remaining 24 patients were positive in their second (n=18), third (n=5) or fourth (n=1) PCR.

CoLab-score

The model obtained through adaptive lasso regression contained 11 variables, which are depicted with their regression coefficients (weights) in table 2.

A larger β-coefficient does not imply that a variable is more important in predicting the odds of testing positive for SARS-CoV-2 since the variables are on different scales. The most important variables are basophils, eosinophils and LD.

As shown in figure 2, the linear predictor clearly discriminates between COVID-19 and non-COVID-19. The linear predictor is converted to CoLab-scores 0–5, with the cut-points depicted in figure 2.

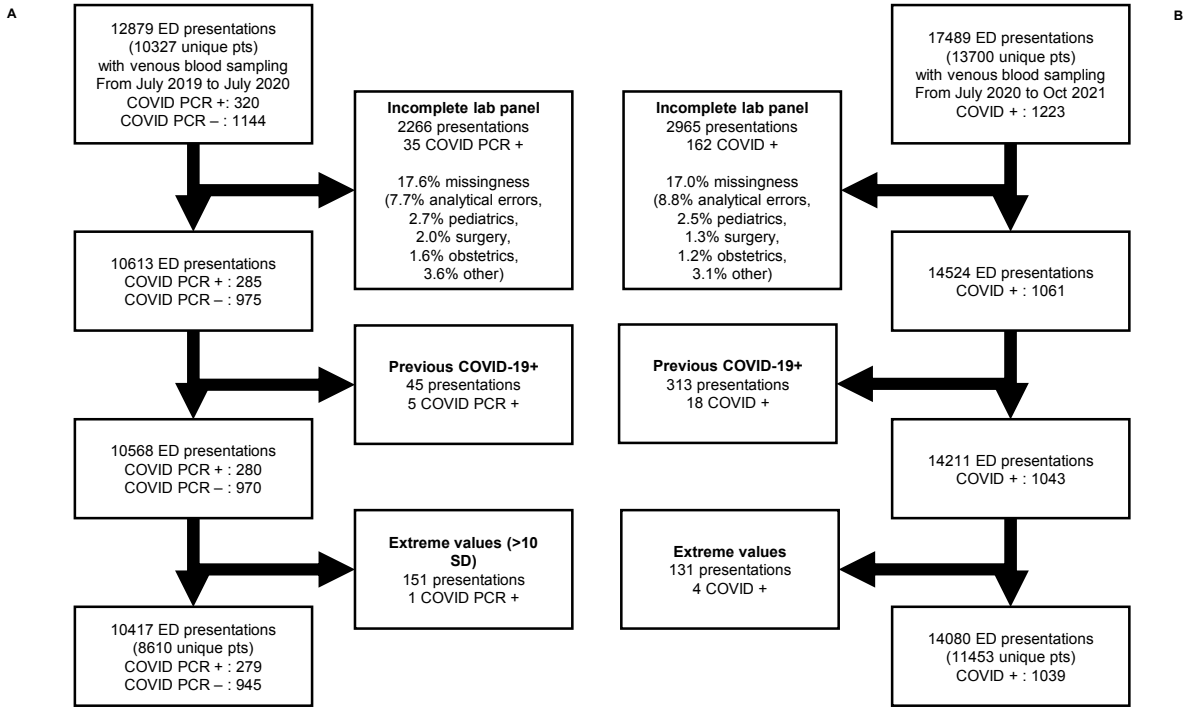

Figure 1 Inclusion flow of patients (pts) in the development (A) and temporal validation (B) data set. All patient admissions with routine venous blood sampling at the emergency department (ED) were included. For the development data set, completeness of the laboratory panel was assessed for all 28 laboratory tests; for the temporal validation data set this was only necessary for 10 laboratory tests. The major causes of missingness are described in the text. In the development data set, presentations with extreme values (>10 SD) were excluded. The same limits were applied to the temporal validation data set (see table 2 for limits).

Table 1 Descriptive statistics of the development data set and laboratory concentrations

	Pre-COVID-19 n=5890	Untested n=3303	PCR negative n=945	PCR positive n=279
Age in years*	61 (21)	60 (21)	66 (18)	69 (15)
Female gender, n (%)	2909 (49.4)	1659 (50.2)	466 (49.3)	95 (34.1)
Specialism, n (%)				
Internal medicine	1648 (28.0)	896 (27.1)	244 (25.8)	71 (25.4)
Surgery	1007 (17.1)	679 (20.6)	51 (5.4)	5 (1.8)
Neurology	775 (13.2)	468 (14.2)	64 (6.8)	5 (1.8)
Pulmonary medicine	714 (12.1)	220 (6.7)	326 (34.5)	167 (59.9)
Cardiology	560 (9.5)	322 (9.7)	145 (15.3)	6 (2.2)
Urology	309 (5.2)	148 (4.5)	15 (1.6)	7 (2.5)
Gastroenterology	306 (5.2)	224 (6.8)	27 (2.9)	1 (0.4)
Geriatrics	189 (3.2)	95 (2.9)	52 (5.5)	15 (5.4)
Orthopaedics	147 (2.5)	109 (3.3)	11 (1.2)	0 (0.0)
Gynaecology	118 (2.0)	82 (2.5)	2 (0.2)	0 (0.0)
Other	117 (2.0)	60 (1.8)	8 (0.8)	2 (0.7)
Haemoglobin* (g/L)	13.2 (2.1)	13.3 (2.0)	13.3 (2.2)	13.8 (1.8)‡
Haematocrit* (L/L)	0.403 (0.059)	0.405 (0.056)	0.405 (0.062)	0.417 (0.047)‡
Erythrocytes* (/pL)	4.41 (0.69)	4.43 (0.66)	4.41 (0.72)	4.61 (0.60)‡
MCV* (fL)	91.8 (6.4)	91.9 (6.1)	92.4 (6.7)	90.7 (5.5)
MCH* (mmol)	1.859 (0.157)	1.876 (0.150)	1.874 (0.172)	1.869 (0.141)
MCHC* (mmol/L)	20.2 (0.9)	20.4 (0.9)	20.3 (1.0)	20.6 (0.8)‡
Thrombocytes* (/nL)	263 (99)	266 (100)	269 (105)	217 (123)‡
Leucocytes† (/nL)	9.30 (7.06–12.16)	8.92 (7.01–11.89)	9.66 (7.17–12.94)	6.33 (4.74–8.48)‡
Neutrophils† (/nL)	6.62 (4.51–9.53)	6.10 (4.42–8.94)	7.01 (4.79–10.02)	4.71 (3.30–6.94)‡
Eosinophils† (/nL)	0.09 (0.03–0.17)	0.09 (0.03–0.18)	0.08 (0.02–0.17)	0.00 (0.00–0.02)‡
Basophils† (/nL)	0.04 (0.02–0.05)	0.04 (0.02–0.05)	0.04 (0.02–0.05)	0.01 (0.01–0.02)‡
Lymphocytes† (/nL)	1.47 (0.93–2.13)	1.56 (1.05–2.18)	1.31 (0.80–2.03)	0.86 (0.59–1.21)‡
Monocytes† (/nL)	0.70 (0.52–0.93)	0.69 (0.52–0.91)	0.74 (0.54–1.01)	0.45 (0.32–0.64)‡
Glucose† (mmol/L)	6.76 (5.83–8.39)	6.68 (5.76–8.14)	6.98 (5.95–8.85)	6.77 (5.98–8.48)‡
Bilirubin† (µmol/L)	7.5 (5.0–11.6)	7.4 (5.1–10.9)	8.3 (5.6–12.4)	8.2 (6.3–11.4)
ASAT† (U/L)	24.0 (19.1–32.2)	26.5 (21.6–35.1)	27.7 (21.7–39.2)	40.7 (30.2–57.2)‡
ALAT† (U/L)	24.3 (17.8–35.3)	25.3 (18.4–36.2)	25.7 (18.4–40.0)	33.7 (23.3–50.0)‡
LD† (U/L)	201 (173–240)	198 (170–236)	215 (178–263)	300 (238–403)‡
CK† (U/L)	82 (51–134)	83 (52–136)	76 (51–125)	124 (62–222)‡
ALP† (IU/L)	83.0 (68.0–105.0)	81.0 (65.8–102.5)	86.9 (67.9–110.0)	71.0 (58.8–85.0)‡
gGT† (U/L)	27.0 (17.0–53.0)	28.4 (18.4–50.5)	37.0 (22.4–68.9)‡	42.0 (28.0–83.5)‡
BUN† (mmol/L)	5.7 (4.3–8.0)	5.8 (4.3–7.8)	6.2 (4.6–9.4)	6.1 (4.7–8.9)
CKD-epi† (mL/min/m^2)	80.9 (58.0–99.1)	85.0 (63.5–103.3)	79.1 (52.1–96.6)	76.6 (54.9–91.2)
Potassium* (mmol/L)	4.06 (0.50)	4.03 (0.49)	4.07 (0.55)	3.91 (0.47)
Sodium* (mmol/L)	139.2 (4.0)	138.5 (3.9)	138.0 (4.3)‡	136.4 (4.1)‡
Chloride* (mmol/L)	104.4 (4.6)	103.8 (4.5)	102.9 (4.8)	101.6 (4.4)‡
Albumin* (g/L)	42.4 (4.9)	42.3 (4.5)	40.8 (4.8)	38.4 (3.8)‡
CRP† (mg/L)	8 (2–41)	5 (1–30)	18 (3–69)‡	77 (37–136)‡

Shown are the laboratory tests routinely requested at ED presentations and their mean/median results (in the development data set) for presentations before the first patient with COVID-19 in the Netherlands ('Pre-COVID-19'), presentations thereafter that were not tested for COVID-19 ('Untested'), tested negative ('PCR negative') and tested positive ('PCR positive').
*Results with normal distribution, where the mean value and SD are shown.
†Results with skewed or heavy tailed distribution, where the median value and IQR are shown.
‡Clinically relevant difference from the pre-COVID-19 category (based on the total allowable error).
ALAT, alanine aminotransferase; ALP, alkaline phosphatase; ASAT, aspartate aminotransferase; BUN, blood urea nitrogen; CK, creatine kinase; CKD-epi, Chronic Kidney Disease Epidemiology Collaboration; CRP, C reactive protein; ED, emergency department; gGT, gamma-glutamyltransferase; LD, lactate dehydrogenase; MCH, mean cellular haemoglobin; MCHC, mean cellular haemoglobin concentration; MCV, mean corpuscular volume.

Table 2 Calculation of the CoLab linear predictor

Variable	β	Exclusion limit	Relative importance (%)
Intercept	−6.885		–
Erythrocytes (/pL)	0.9379	Erythrocytes <2.9/pL	52
Leucocytes (/nL)	−0.1298		46
Eosinophils (/nL)	−6.834		86
Basophils (/nL)	−47.70	Basophils >0.33/nL	100
Log_{10} of bilirubin (µmol/L)	−1.142	Bilirubin >169 µmol/L	26
Log_{10} of LD (U/L)	5.369	LD >1564 U/L	58
Log_{10} of ALP (IU/L)	−3.114	AF >1000 IU/L	45
Log_{10} of gGT (U/L)	0.3605	gGT >1611 U/L	11
Albumin (g/L)	−0.1156		45
CRP (mg/L)	0.002560		15
Age (years)	0.002275		4

The CoLab linear predictor (LP) is calculated by summing the intercept and the products of the 11 variables with their corresponding coefficients (βs). CoLab LP=−6.885+[erythrocytes]×0.9379−[leucocytes]×0.1298−[eosinophils]×6.834−[basophils]×47.7−log_{10}[bilirubin]×1.142+log_{10}[LD]×5.369−log_{10}[ALP]×3.114+log_{10}[gGT]×0.3605−[albumin]×0.1156+[CRP]×0.02560+[age]×0.002275. The LP can be converted into a CoLab-score (see figure 2) or into a probability if the prevalence is known or estimated (see details in online supplemental material 1). The CoLab-score is not valid if any of the variables exceeds the limits in the third column. The relative importance ranks the importance of variables in predicting the outcome, relative to the most important variable (in this case basophils).
ALP, alkaline phosphatase; CRP, C reactive protein; gGT, gamma-glutamyltransferase; LD, lactate dehydrogenase.

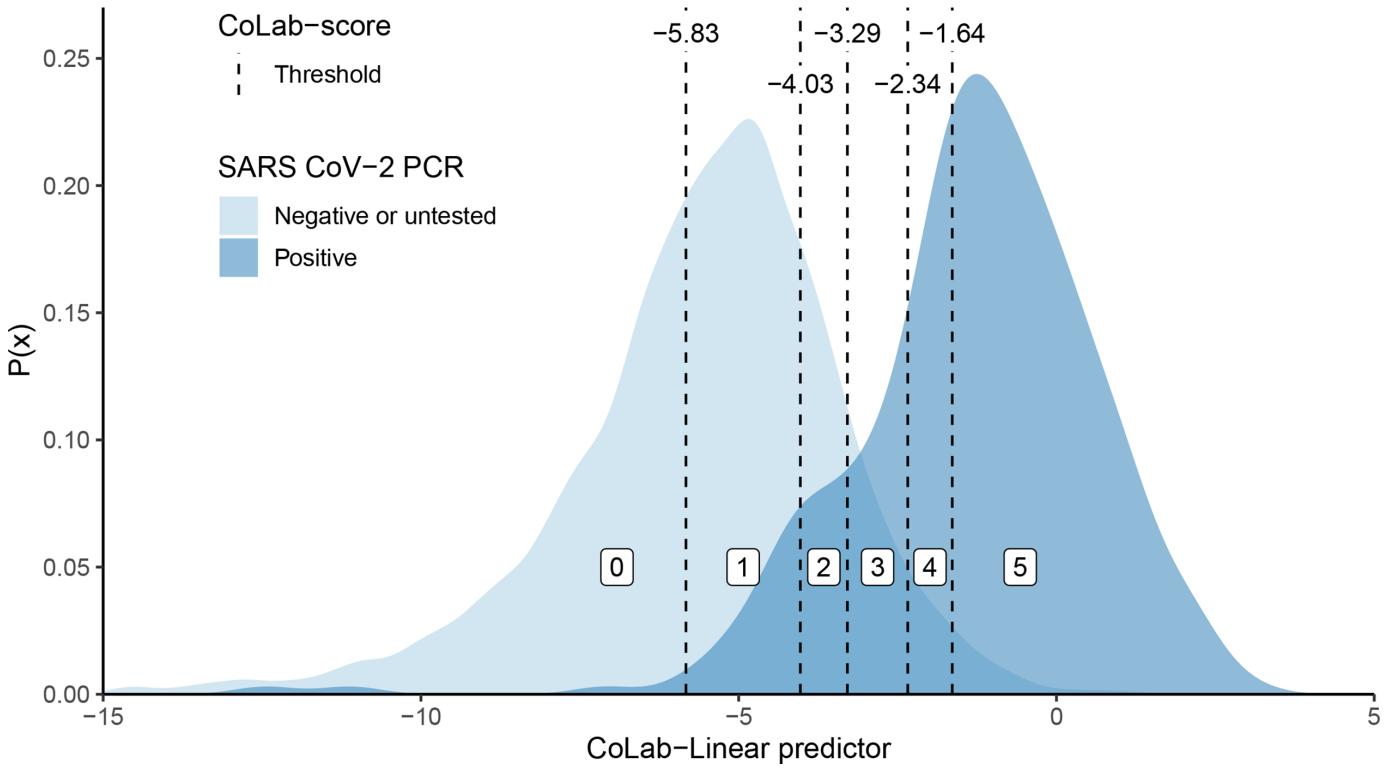

Figure 2 Probability density plot of the CoLab linear predictor. The probability density plots for patients with COVID-19 (dark blue) and those without COVID-19 (light blue) are plotted against the linear predictor (see table 2). The CoLab-score cut-offs (−5.83, −4.02, −3.29, −2.34 and −1.64) are depicted with vertical dashed lines. The white-boxed numbers (between the cut-offs) represent the corresponding CoLab-score. Note that while the area under both curves is identical (since these are probability density functions), in absolute numbers the 'negative or untested' group is about 36 times larger than the PCR-positive group.

Internal validation

The model was validated in the period starting from the first COVID-19 infection to July 2020, and in this period the mean prevalence was 7.2%. The AUC of the CoLab-score is 0.930 (95% CI 0.909 to 0.945).

Diagnostic performance is shown in table 3. A CoLab-score of 0 has an NPV of 0.997 (95% CI 0.993 to 0.999) and a PPV of 0.115 (95% CI 0.0934 to 0.147); one-third (38%, 95% CI 28% to 514%) of all ED presentations were assigned this score and can therefore be safely excluded. Conversely, 6% (95% CI 6% to 8%) of the ED patients had a CoLab-score of 5. Given the PPV of this score (0.683, 95% CI 0.628 to 0.746; NPV: 0.970, 95% CI 0.963 to 0.978), subsequent PCR testing is advised.

Temporal validation

As the CoLab-score was developed at our centre after the first COVID-19 wave in the Netherlands, the performance was evaluated at our centre from July 2020 until October 2021. Laboratory results from 17 489 ED presentations were collected. After applying the inclusion flow as shown in figure 1B and 14080 presentations remained, of which 1039 were associated with a COVID-19 infection.[19]

The mean prevalence in this period was 7.4%. The AUC of the CoLab-score in the temporal validation set is 0.916 (95% CI 0.906 to 0.927). The performance is comparable with the development cohort, although sensitivity is slightly lower and specificity slightly higher (cf, table 3 and table 4). The temporal validation data set was also split into three phases according to the dominant SARS-CoV-2 variants and vaccine roll-out (see figure 1 in online supplemental material 2). The discriminative ability was not lower in the second or third phase compared with the first phase. Diagnostic performance is preserved in terms of sensitivity and specificity, except a moderately reduced sensitivity of scores ≥3 in the third phase as compared with the first phase. PPV and NPV are incomparable due to different prevalence/pretest probabilities in each phase (see table 1 in online supplemental material 2).

In terms of the predicted probabilities, model calibration shows that overall predicted probabilities are too low (see figure 1 in online supplemental material 3 for the calibration plot), which is expected since the prevalence differs and the intercept has to be adjusted to the prevalence.

In this period at least 22 COVID-19-positive patients were identified by the CoLab-score, who initially did not present with COVID-19-specific symptoms. Most patients had neurological or orthopaedic presenting symptoms.

External validation

For external validation, data obtained from three other centres were used: centre 1 (n=1284, 52 COVID-19-positive), centre 2 (n=2899, 99 COVID-19-positive) and centre 3 (n=3545, 336 COVID-19-positive).[19] The inclusion flow is summarised in figure 3. The COVID-19 prevalence differed among the three centres (4.0%, 3.4% and 9.5%, respectively) and was lower in centres 1 and 2

Table 3 Bootstrapped diagnostic performance of the CoLab-score in the development data set

CoLab-score	Sensitivity	Specificity	PPV	NPV	TP	TN	FP	FN	% of population
0	0.984 (0.969 to 0.991)	0.410 (0.302 to 0.543)	0.115 (0.094 to 0.147)	0.997 (0.993 to 0.999)	273.4 (241.2 to 304.4)	1470.9 (1081.1 to 1950.9)	2119.1 (1633.5 to 2507.6)	4.6 (2.6 to 8.6)	38.0 (28.0 to 51.0)
≤1	0.912 (0.892 to 0.952)	0.785 (0.741 to 0.827)	0.248 (0.207 to 0.300)	0.991 (0.989 to 0.995)	253.5 (226.5 to 287.0)	2817.1 (2655.4 to 2961.2)	772.9 (623.2 to 934.5)	24.5 (13.4 to 30.2)	73.3 (69.3 to 77.3)
≤2	0.856 (0.816 to 0.895)	0.880 (0.864 to 0.900)	0.357 (0.315 to 0.415)	0.988 (0.984 to 0.991)	238.1 (209.6 to 267.9)	3160.8 (3100.7 to 3233.7)	429.1 (357.3 to 487.1)	39.9 (28.5 to 52.4)	82.9 (80.9 to 83.9)
≤3	0.757 (0.706 to 0.809)	0.951 (0.944 to 0.959)	0.546 (0.496 to 0.604)	0.981 (0.976 to 0.985)	210.4 (183.4 to 240.2)	3415.1 (3378.0 to 3456.4)	174.9 (147.0 to 199.3)	67.6 (51.9 to 84.9)	90.0 (89.0 to 91.0)
≤4	0.612 (0.530 to 0.706)	0.978 (0.972 to 0.983)	0.683 (0.628 to 0.746)	0.970 (0.963 to 0.978)	170.2 (141.6 to 204.9)	3510.6 (3476.8 to 3547.5)	79.4 (60.3 to 100.4)	107.9 (79.1 to 134.0)	93.7 (91.7 to 93.7)

The development data set was internally validated for the period March 2020–July 2020 (n=3868). The optimism-adjusted bootstrapped sensitivity, specificity, PPV, NPV, TP, TN, FP and FN and the fraction of presentations (%) are shown for fixed cut-offs (CoLab-score 0 to ≤4). The numbers in brackets represent the 95% optimism-adjusted bootstrapped CI. The first column defines the threshold above which CoLab-score a patient is considered positive. Note that '0' lists the sensitivity and NPV of CoLab-score 0 and '≤4' lists the specificity and PPV of CoLab-score 5. Also note that TP, TN, FP and FN are not whole numbers as these are obtained through bootstrapping and each bootstrap replicate contains a different number of controls and cases.
FN, false negative; FP, false positive; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

Table 4 Diagnostic performance of the CoLab-score in the validation data set (temporal) and three external hospitals

CoLab-score	Validation set	Sensitivity	Specificity	PPV	NPV	TP	TN	FP	FN
0	Temporal	0.967 (0.956 to 0.978)	0.420 (0.411 to 0.428)	0.117 (0.115 to 0.119)	0.994 (0.992 to 0.996)	1005 (993 to 1016)	5476 (5366 to 5587)	7565 (7454 to 7675)	34 (23 to 46)
	Centre 1	1.000 (1.000 to 1.000)	0.331 (0.307 to 0.358)	0.059 (0.057 to 0.061)	1.000 (1.000 to 1.000)	52 (52 to 52)	410 (380 to 443)	827 (794 to 857)	0 (0 to 0)
	Centre 2	0.961 (0.922 to 0.990)	0.351 (0.333 to 0.369)	0.052 (0.049 to 0.054)	0.996 (0.992 to 0.999)	99 (95 to 102)	985 (935 to 1035)	1823 (1773 to 1873)	4 (1 to 8)
	Centre 3	0.970 (0.950 to 0.988)	0.322 (0.306 to 0.338)	0.130 (0.126 to 0.133)	0.991 (0.984 to 0.996)	327 (320 to 333)	1042 (991 to 1092)	2193 (2143 to 2244)	10 (4 to 17)
≤1	Temporal	0.888 (0.870 to 0.908)	0.791 (0.783 to 0.798)	0.253 (0.245 to 0.261)	0.989 (0.987 to 0.991)	923 (904 to 943)	10 311 (10 215 to 10 401)	2730 (2640 to 2826)	116 (96 to 135)
	Centre 1	0.923 (0.846 to 0.981)	0.694 (0.669 to 0.720)	0.113 (0.101 to 0.124)	0.995 (0.991 to 0.999)	48 (44 to 51)	858 (828 to 891)	379 (346 to 409)	4 (1 to 8)
	Centre 2	0.913 (0.854 to 0.961)	0.678 (0.661 to 0.696)	0.094 (0.087 to 0.101)	0.995 (0.992 to 0.998)	94 (88 to 99)	1905 (1857 to 1953)	903 (855 to 951)	9 (4 to 15)
	Centre 3	0.914 (0.881 to 0.944)	0.674 (0.657 to 0.691)	0.226 (0.216 to 0.236)	0.987 (0.982 to 0.991)	308 (297 to 318)	2180 (2126 to 2234)	1055 (1001 to 1109)	29 (19 to 40)
≤2	Temporal	0.820 (0.796 to 0.843)	0.894 (0.889 to 0.899)	0.382 (0.367 to 0.396)	0.984 (0.982 to 0.986)	852 (827 to 876)	11 661 (11 591 to 11 729)	1380 (1312 to 1450)	187 (163 to 212)
	Centre 1	0.808 (0.692 to 0.904)	0.811 (0.788 to 0.832)	0.152 (0.129 to 0.176)	0.990 (0.984 to 0.995)	42 (36 to 47)	1003 (975 to 1029)	234 (208 to 262)	10 (5 to 16)
	Centre 2	0.845 (0.777 to 0.913)	0.801 (0.785 to 0.815)	0.135 (0.122 to 0.147)	0.993 (0.990 to 0.996)	87 (80 to 94)	2248 (2205 to 2289)	560 (519 to 603)	16 (9 to 23)
	Centre 3	0.890 (0.855 to 0.923)	0.794 (0.779 to 0.808)	0.311 (0.294 to 0.328)	0.986 (0.981 to 0.990)	300 (288 to 311)	2569 (2521 to 2615)	666 (620 to 714)	37 (26 to 49)
≤3	Temporal	0.710 (0.682 to 0.738)	0.962 (0.958 to 0.965)	0.596 (0.573 to 0.618)	0.977 (0.974 to 0.979)	738 (709 to 767)	12 540 (12 496 to 12 582)	501 (459 to 545)	301 (272 to 330)
	Centre 1	0.750 (0.635 to 0.865)	0.909 (0.892 to 0.925)	0.257 (0.213 to 0.306)	0.989 (0.983 to 0.994)	39 (33 to 45)	1124 (1104 to 1144)	113 (93 to 133)	13 (7 to 19)
	Centre 2	0.660 (0.563 to 0.748)	0.897 (0.885 to 0.908)	0.190 (0.163 to 0.218)	0.986 (0.983 to 0.990)	68 (58 to 77)	2519 (2486 to 2549)	289 (259 to 322)	35 (26 to 45)
	Centre 3	0.766 (0.718 to 0.810)	0.887 (0.876 to 0.898)	0.413 (0.386 to 0.442)	0.973 (0.968 to 0.978)	258 (242 to 273)	2869 (2835 to 2905)	366 (330 to 400)	79 (64 to 95)
≤4	Temporal	0.585 (0.556 to 0.615)	0.984 (0.982 to 0.987)	0.750 (0.724 to 0.778)	0.968 (0.965 to 0.970)	608 (578 to 639)	12 838 (12 811 to 12 866)	203 (175 to 230)	431 (400 to 461)
	Centre 1	0.654 (0.519 to 0.788)	0.951 (0.939 to 0.962)	0.359 (0.293 to 0.435)	0.985 (0.979 to 0.991)	34 (27 to 41)	1176 (1161 to 1190)	61 (47 to 76)	18 (11 to 25)
	Centre 2	0.534 (0.437 to 0.621)	0.952 (0.943 to 0.959)	0.287 (0.239 to 0.339)	0.982 (0.979 to 0.986)	55 (45 to 64)	2672 (2649 to 2693)	136 (115 to 159)	48 (39 to 58)
	Centre 3	0.665 (0.611 to 0.718)	0.930 (0.921 to 0.938)	0.497 (0.462 to 0.534)	0.964 (0.958 to 0.969)	224 (206 to 242)	3008 (2980 to 3036)	227 (199 to 255)	113 (95 to 131)

Sensitivity, specificity, PPV, NPV, TP, TN, FP and FN are shown for fixed cut-offs (CoLab-score 0 to ≤4) with bootstrapped 95% CI in parentheses. Note that '0' lists the sensitivity and NPV of CoLab-score 0 and '≤4' lists the specificity and PPV of CoLab-score 5.
FN, false negative; FP, false positive; NPV, negative predictive value; PPV, positive predictive value; TN, true negative; TP, true positive.

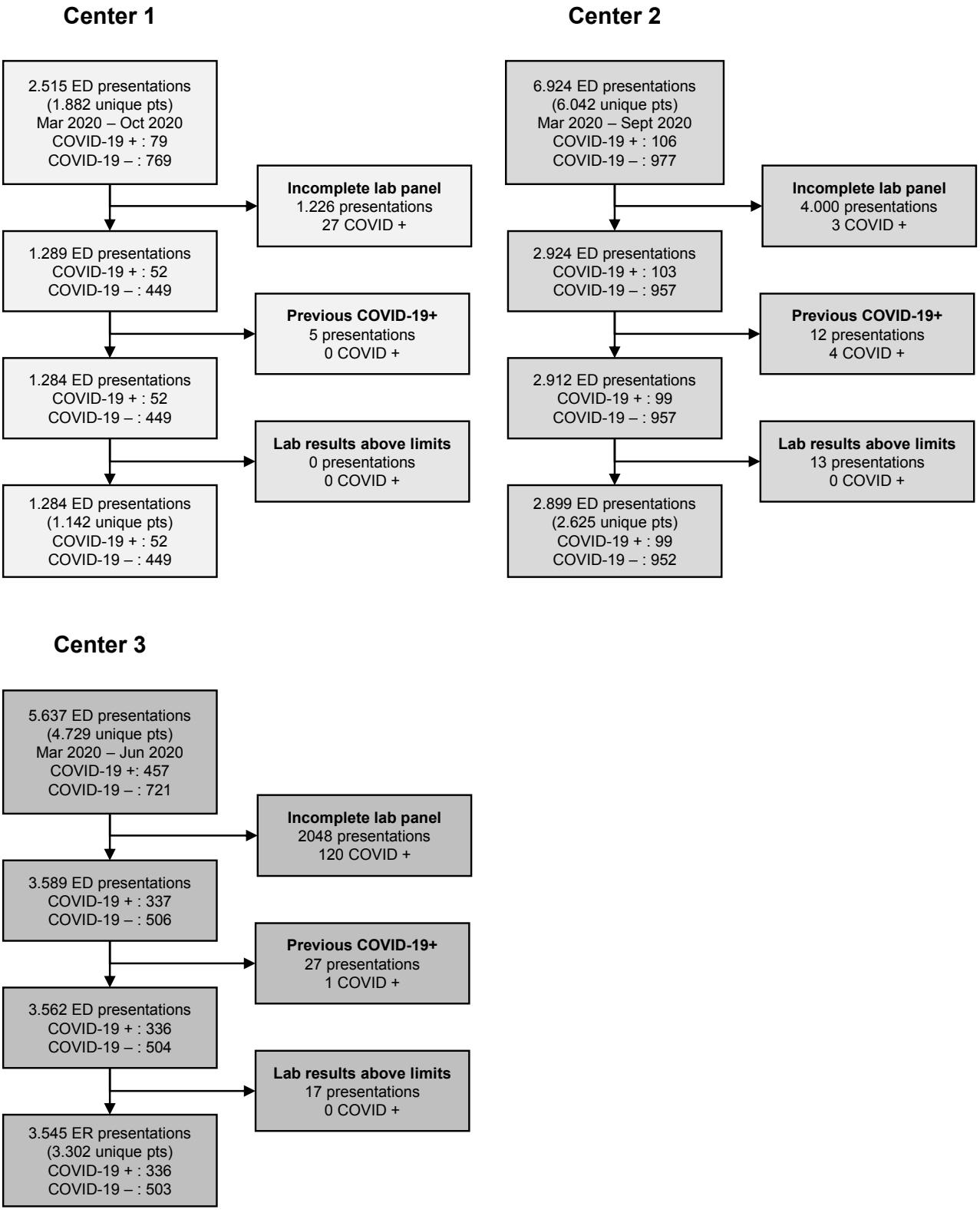

Figure 3 Inclusion flow of emergency department (ED) patients (pts) in three external centres. All ED presentations with routine venous blood sampling were included. Missingness of laboratory panels was assessed for the 11 variables in the CoLab-score (see table 2). Re-presentations after a positive PCR result or clinical COVID-19 registration were excluded as 'previous COVID-19+'. Presentations with any laboratory result above the limits of the CoLab-score (see table 2) were excluded.

and higher in centre 3 than in the development data set. The AUCs of the CoLab-score are 0.904 (95% CI 0.866 to 0.942), 0.886 (95% CI 0.851 to 0.922) and 0.891 (95% CI 0.872 to 0.909) for centres 1, 2 and 3, respectively.

Diagnostic performance is shown in table 4. The sensitivity of CoLab-score 0 in all centres is ≥0.96. Therefore, the NPV of CoLab-score 0 was more than 99%. Calibration plots for external centres are shown in figure 1 in online supplemental material 3. The observed fraction of COVID-19 positives is slightly lower than expected in centres 1 and 2. For centre 3, low probabilities appear slightly underestimated and high probabilities slightly overestimated.

DISCUSSION

Given the impact of COVID-19 on society and healthcare, there is a need for simple and fast detection of patients with a possible COVID-19 infection in the ED. The CoLab-score described in this study is a fast and accurate risk score to triage patients presenting to the ED based on 10 routine blood biomarkers and age.

The main strength of this study is that this score can be used as an early warning or triaging tool for the ED population presenting with abdominal pain, chest pain, shortness of breath, syncope, sepsis or other non-specific complaints where a routine blood panel is requested. This is in contrast to the vast majority of COVID-19 diagnostic models that have been developed on a preselected population of PCR-tested patients.[9 21–27] Moreover, the CoLab-score requires only routine blood tests, instead of (features from) imaging such as CT scans or laboratory tests that are not routinely collected in the ED, for example, interleukin 6 or 3-hydroxybutyric acid.[4] Compared with lateral flow tests (LFTs), which provide a dichotomous result within 30 min and are widely adopted in EDs, the CoLab-score is a continuous score. The lowest CoLab-scores (0–1) offer higher sensitivity and are therefore more suitable to rule out COVID-19 than LFT, which is only moderately sensitive (although more specific).[28 29]

Two other studies have been published which are similar to this study.[9 10] Interestingly, the study by Soltan *et al*[10] ranked basophils and eosinophils as the two most important features in predicting the outcome, similar to our results. Eosinophils were also seen as one of the most important features by Plante *et al*.[9] However, both studies focus on an artificial intelligence/machine learning approach. While their approach likely results in higher predictive performance, due to the ability of machine learning models to capture non-linear and interaction effects, the goal of this study was to develop a simple, fast and robust model that can easily be implemented in current hospital information systems.

Since this is a retrospective case–control study, there are some unavoidable missing data. In our cohort 17.6% of the ED presentations could not be used due to one or more missing laboratory results. This is lower or equal to similar studies: 22%,[24] 17%[22] and 11%.[27] Important to note is that 7.7% of missingness is due to analytical errors, which can be assumed to be missing completely at random. For the remaining 9.9% of missingness, the full laboratory panel was most frequently missing for paediatric, obstetric and surgery patients. These patients are presenting with specific complaints for which specific laboratory tests are requested and hence do not match the inclusion criteria for a routine blood panel. Overall the missingness was significantly lower in the PCR-tested group versus the untested group (χ^2 test p<0.001). It is assumed that all presentations in the untested group are COVID-19-negative. However, some presentations with asymptomatic COVID-19 could be present in the untested control group. The impact of these 'false controls' is most likely small as other studies indicate that there is a very low positivity rate among asymptomatic ED presentations (only a few in over 1000 tested asymptomatic cases).[30 31] The vast majority of controls were not tested for COVID-19 because they were either prepandemic or untested patients (89% in the development data set). Clinical data always contain some unavoidable 'noise' in the form of misregistrations, misdiagnoses or patients who were missed. We have tried to mitigate this by including a large prepandemic control group and including all PCR tests within 1 week after discharge.

In the external centres, there is a high level of missingness as a result of an incomplete laboratory panel. In the case of centres 1 and 2, only internal medicine ED presentations were tested with a laboratory panel containing the 10 tests required for the CoLab-score. The ED laboratory panel of other disciplines (eg, urology, surgery or paediatrics) differed and did not contain the required tests. Nevertheless, the majority of patients with COVID-19 were internal medicine ED presentations, reflected by the few PCR-positive patients excluded. Due to these high levels of missingness, the results of the external centres cannot be used to show that the CoLab-score generalises to the entire ED population. Rather, the results show that for the majority of COVID-19-positive patients presenting to the ED, a routine laboratory panel is available from which the CoLab-score can be calculated and that the performance of the CoLab-score in this population is comparable with the development population. Differences in the distribution of CoLab variables between centres are shown in figure 2 in online supplemental material 3.

The performance of the CoLab-score is affected by the time between the onset of symptoms and ED presentations. The score increases with the duration of symptoms and gradually decreases after day 7 (see figure 1 in online supplemental material 4 for a plot of the duration of COVID-19-related symptoms and the CoLab linear predictor). As a consequence, some patients with COVID-19 with early or late presentation after onset of symptoms can be missed. Optimal performance of the CoLab-score is achieved when the onset of symptoms is >1 and <10 days prior to ED presentation. Chemotherapy that causes myeloid suppression will decrease neutrophilic, basophilic and eosinophilic counts and thereby

'falsely' increasing the CoLab-score. Conversely, patients with COVID-19 with severe anaemia could have 'falsely' lowered CoLab-scores. To minimise false negatives, we have therefore advised to report CoLab-scores only when the concentration of erythrocytes is ≥2.9/pL.

It was chosen to exclude re-presentations after a previous presentation with COVID-19. Since the median time between initial presentation and re-presentation was 12 days, these patients were most likely not reinfected patients, but patients who deteriorated after initial presentation/treatment. Given that the CoLab-score follows the host immune response, the score is time-sensitive (see figure 1 in online supplemental material 4). Including these patients would impact the performance of the CoLab-score as patients in a later phase of the disease show different biomarker profiles. The CoLab-score is aimed towards alerting clinicians to patients presenting with a novel SARS-CoV-2 infection, rather than patients who deteriorate after treatment for COVID-19. Other re-presentations were not excluded, which results in some patients appearing multiple times in a data set. This was not adjusted for in the regression model since the assumption was made that ED presentations are independent observations. The median time between re-presentations is 38 days, most likely resulting in variations in laboratory results between presentations and hence little to no correlation between presentations. A sensitivity analysis was performed whereby only the first presentation was included for each patient (table 1 in online supplemental material 4) but no difference was found in performance in terms of sensitivity, specificity and AUC.

The CoLab-score does not serve as a replacement for PCR testing or LFT and can be used to guide PCR testing when routine blood tests are available. Important to note is that the CoLab-score is only valid for ED presentations where routine blood testing is requested, and as a consequence does not generalise to the ED population who is otherwise well and does not undergo routine blood testing. Using the CoLab-score in a symptomatic/PCR-tested cohort also results in different diagnostic performance characteristics, as compared with using the score on the full ED cohort (see table 1 in online supplemental material 4).

Finally, the CoLab-score could lead to false positives by other viral infections. However, in a historical patient cohort, the CoLab-score had only limited discriminative ability in separating influenza-PCR-negative from influenza-PCR-positive patients (see figure 2 in online supplemental material 4), implying specificity for SARS-CoV-2. Since the CoLab-score reflects the host response to the virus, it is hypothesised that the CoLab-score could also be sensitive to future SARS-CoV-2 variants. This is supported by the fact that the discriminative ability is sustained in periods with different dominant variants, although the sensitivity of scores ≥3 is somewhat lower in the third phase (see table 1 in online supplemental material 2). Although vaccination status is not registered for all presenting patients, in a small subgroup of 12 patients for whom vaccination status was registered and were COVID-19-positive, 8 of 12 patients had the highest CoLab-score (score=5) (see figure 2 in online supplemental material 2). Continuous assessment of the performance of the CoLab-score is required due to the emergence of new variants and changes in the host's immune response.

To conclude, the CoLab-score developed and validated in this study, based on 10 routine laboratory results and age, is available within 1 hour for any patient presenting to the ED where routine blood testing is requested. The score can be used by clinicians to guide PCR testing or triage patients and helps to identify COVID-19 in patients presenting to the ED with abdominal pain, chest pain, shortness of breath, syncope, sepsis or other non-specific complaints where a routine blood panel is requested. The lowest CoLab-score can be used to effectively rule out a possible SARS-CoV-2 infection, the highest score to alert physicians to a possible infection. The CoLab-score is therefore a valuable tool to rule out COVID-19, guide PCR testing and is available to any centre with access to routine laboratory tests.

Author affiliations
[1]Department of Laboratory Medicine, Catharina Hospital, Eindhoven, The Netherlands
[2]Faculty of Biomedical Engineering, Eindhoven University of Technology, Eindhoven, The Netherlands
[3]Department of Emergency Medicine, Catharina Hospital, Eindhoven, The Netherlands
[4]Department of Internal Medicine, Catharina Hospital, Eindhoven, The Netherlands
[5]Department of Pulmonary Diseases, Catharina Hospital, Eindhoven, The Netherlands
[6]Department of Emergency Medicine, Catharina Ziekenhuis, Eindhoven, The Netherlands
[7]Department of Clinical Chemistry and Hematology, Zuyderland Medical Centre Heerlen, Heerlen, The Netherlands
[8]Atalmedial Diagnostic Centers, Amsterdam, The Netherlands
[9]Department of Medical Microbiology and Infection Prevention, Alrijne Hospital, Leiderdorp, The Netherlands
[10]Department of Clinical Chemistry, Gelre Hospitals, Apeldoorn, The Netherlands
[11]Amsterdam University Medical Center, University of Amsterdam, Amsterdam, The Netherlands

Contributors A-KB: guarantor, conceptualisation (lead), data curation (lead), funding acquisition (lead), investigation (equal), methodology (equal), supervision (equal), writing - original draft (equal), writing - review and editing (equal). RD: data curation (equal), formal analysis (equal), investigation (equal), methodology (lead), software (lead), visualisation (lead), writing - original draft (equal), writing - review and editing (equal). MMa, RHHvB, SB: conceptualisation (supporting), resources (supporting), supervision (supporting), validation (supporting), writing - review and editing (equal). HSMA: conceptualisation (supporting), resources (supporting), supervision (supporting), validation (equal), writing - review and editing (equal). WAHMT: conceptualisation (supporting), resources (supporting), supervision (supporting), validation (supporting), writing - review and editing (equal). ML, RJHM, MMB, JJK, MMe: resources (equal), validation (equal), writing - review and editing (equal). JJvS: resources (supporting), validation (supporting), writing - review and editing (equal). NAWvR: methodology (supporting), resources (supporting), supervision (equal), writing - review and editing (equal). VS: conceptualisation (equal), funding acquisition (equal), project administration (lead), resources (equal), supervision (lead), writing - review and editing (equal).

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not required.

Ethics approval The study was reviewed by the Medical Research Ethics Committees United (MEC-U) under study number W20.071, which confirmed that the Medical Research Involving Human Subjects Act (in Dutch: WMO) does not apply to this study. The study was thereafter reviewed and approved by the internal hospital review board.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available in a public, open access repository. Data is de-indentified.

ORCID iDs
Ruben Deneer http://orcid.org/0000-0002-6011-7783
Mathie Leers http://orcid.org/0000-0001-5186-5600

REFERENCES

1 Disease C. (COVID-19) situation reports. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/ [Accessed 4 Feb 2021].
2 Guan W-jie, Ni Z-yi, Hu Y, *et al*. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med Overseas Ed* 2020;382:1708–20.
3 Vetter P, Vu DL, L'Huillier AG, *et al*. Clinical features of covid-19. *BMJ* 2020;369:m1470.
4 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.
5 Albahri AS, Hamid RA, Alwan JK, *et al*. Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (COVID-19): a systematic review. *J Med Syst* 2020;44:122.
6 Hooli S, King C. Generalizability of coronavirus disease 2019 (COVID-19) clinical prediction models. *Clin Infect Dis* 2020;71:897.
7 Prokop M, van Everdingen W, van Rees Vellinga T, *et al*. CO-RADS: a categorical CT assessment scheme for patients suspected of having COVID-19-Definition and evaluation. *Radiology* 2020;296:E97–104.
8 Goodacre S, Thomas B, Sutton L, *et al*. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: the Priest observational cohort study. *PLoS One* 2021;16:e0245840.
9 Plante TB, Blau AM, Berg AN, *et al*. Development and external validation of a machine learning tool to rule out COVID-19 among adults in the emergency department using routine blood tests: a large, multicenter, real-world study. *J Med Internet Res* 2020;22:e24048.
10 Soltan AAS, Kouchaki S, Zhu T, *et al*. Rapid triage for COVID-19 using routine clinical data for patients attending Hospital: development and prospective validation of an artificial intelligence screening test. *Lancet Digit Health* 2021;3:e78–87.
11 R Core Team. R: a language and environment for statistical computing, 2020. Available: https://www.r-project.org/
12 Zou H. The adaptive LASSO and its oracle properties. *J Am Stat Assoc* 2006;101:1418–29.
13 Tibshirani R. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society: Series B* 1996;58:267–88.
14 Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.
15 King G, Zeng L. Logistic regression in rare events data. *Polit. anal.* 2001;9:137–63.
16 Efron B. Estimating the error rate of a prediction rule: improvement on cross-validation. *J Am Stat Assoc* 1983;78:316–31.
17 Robin X, Turck N, Hainard A, *et al*. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics* 2011;12:77.
18 Harrell Jr FE. rms: regression modeling strategies, 2021. Available: https://cran.r-project.org/package=rms
19 Boer A-K, Deneer R. Source data for: development and validation of an early warning score to identify COVID-19 in the emergency department based on routine laboratory tests: a multicenter case-control study. *Dryad Digit Repos* 2021.
20 Ricós C, Alvarez V, Cava F, *et al*. Current databases on biological variation: pros, cons and progress. *Scand J Clin Lab Invest* 1999;59:491–500.
21 Brinati D, Campagner A, Ferrari D, *et al*. Detection of COVID-19 infection from routine blood exams with machine learning: a feasibility study. *J Med Syst* 2020;44:1–12.
22 Joshi RP, Pejaver V, Hammarlund NE, *et al*. A predictive tool for identification of SARS-CoV-2 PCR-negative emergency department patients using routine test results. *J Clin Virol* 2020;129:104502.
23 Qin L, Yang Y, Cao Q, *et al*. A predictive model and scoring system combining clinical and CT characteristics for the diagnosis of COVID-19. *Eur Radiol* 2020;30:6797–807.
24 Kurstjens S, van der Horst A, Herpers R, *et al*. Rapid identification of SARS-CoV-2-infected patients at the emergency department using routine testing. *Clin Chem Lab Med* 2020;58:1587–93.
25 Fink DL, Khan PY, Goldman N, *et al*. Development and internal validation of a diagnostic prediction model for COVID-19 at time of admission to hospital. *QJM An Int J Med* 2021;114:699–705.
26 Giamello JD, Paglietta G, Cavalot G, *et al*. A simple tool to help ruling-out Covid-19 in the emergency department: derivation and validation of the LDH-CRP-Lymphocyte (LCL) score. *Emerg Care J* 2020;16.
27 Tordjman M, Mekki A, Mali RD, *et al*. Pre-Test probability for SARS-Cov-2-related infection score: the Paris score. *PLoS One* 2020;15:e0243342.
28 Peto T, Affron D, Afrough B, UK COVID-19 Lateral Flow Oversight Team. COVID-19: rapid antigen detection for SARS-CoV-2 by lateral flow assay: a national systematic evaluation of sensitivity and specificity for mass-testing. *EClinicalMedicine* 2021;36:100924.
29 García-Fiñana M, Hughes DM, Cheyne CP, *et al*. Performance of the Innova SARS-CoV-2 antigen rapid lateral flow test in the Liverpool asymptomatic testing pilot: population based cohort study. *BMJ* 2021;374:n1637.
30 Ford JS, Parikh A, Sandhu R, *et al*. Testing asymptomatic emergency department patients for coronavirus disease 2019 (COVID-19) in a low-prevalence region. *Acad Emerg Med* 2020;27:771–4.
31 Ravani P, Saxinger L, Chandran U, *et al*. COVID-19 screening of asymptomatic patients admitted through emergency departments in Alberta: a prospective quality-improvement study. *CMAJ Open* 2020;8:E887–94.