BMJ Open

COVID outcome prediction in the emergency department (COPE): using retrospective Dutch hospital data to develop simple and valid models for predicting mortality and need for intensive care unit admission in patients who present at the emergency department with suspected COVID-19

David van Klaveren ![ORCID],[1,2] Alexandros Rekkas,[3] Jelmer Alsma,[4]
Rob J C G Verdonschot,[5] Dick T J J Koning,[6] Marlijn J A Kamps,[6] Tom Dormans,[7]
Robert Stassen,[8] Sebastiaan Weijer,[9] Klaas-Sierk Arnold,[10] Benjamin Tomlow,[11]
Hilde R H de Geus,[12] Rozemarijn L van Bruchem-Visser ![ORCID],[4] Jelle R Miedema,[13]
Annelies Verbon,[14] Els van Nood,[15] David M Kent,[2] Stephanie C E Schuit,[16]
Hester Lingsma[1]

![Check for updates]

ABSTRACT

Objectives Develop simple and valid models for predicting mortality and need for intensive care unit (ICU) admission in patients who present at the emergency department (ED) with suspected COVID-19.

Design Retrospective.

Setting Secondary care in four large Dutch hospitals.

Participants Patients who presented at the ED and were admitted to hospital with suspected COVID-19. We used 5831 first-wave patients who presented between March and August 2020 for model development and 3252 second-wave patients who presented between September and December 2020 for model validation.

Outcome measures We developed separate logistic regression models for in-hospital death and for need for ICU admission, both within 28 days after hospital admission. Based on prior literature, we considered quickly and objectively obtainable patient characteristics, vital parameters and blood test values as predictors. We assessed model performance by the area under the receiver operating characteristic curve (AUC) and by calibration plots.

Results Of 5831 first-wave patients, 629 (10.8%) died within 28 days after admission. ICU admission was fully recorded for 2633 first-wave patients in 2 hospitals, with 214 (8.1%) ICU admissions within 28 days. A simple model—COVID outcome prediction in the emergency department (COPE)—with age, respiratory rate, C reactive protein, lactate dehydrogenase, albumin and urea captured most of the ability to predict death. COPE was well calibrated and showed good discrimination for mortality in second-wave patients (AUC in four hospitals: 0.82 (95%

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ We used a large representative sample of patients who presented at the emergency department with suspected COVID-19 and were admitted to hospital in the first wave of the pandemic (n=5831) to develop a tool for predicting death and need for intensive care unit (ICU) admission ('COVID outcome prediction in the emergency Department' (COPE)).

⇒ COPE is based on quickly and objectively obtainable predictors, in contrast with previously proposed models.

⇒ COPE discriminated well and was well-calibrated in temporal and geographical validation in second wave patients of the pandemic (n=3235).

⇒ For patients above the age of 70, the need for ICU admission probably differed from observed decisions on ICU admission and required extrapolation of the age effect.

⇒ External validation of COPE in hospitals that were not present in the development data is required.

CI 0.78 to 0.86); 0.82 (95% CI 0.74 to 0.90); 0.79 (95% CI 0.70 to 0.88); 0.83 (95% CI 0.79 to 0.86)). COPE was also able to identify patients at high risk of needing ICU admission in second-wave patients (AUC in two hospitals: 0.84 (95% CI 0.78 to 0.90); 0.81 (95% CI 0.66 to 0.95)).

Conclusions COPE is a simple tool that is well able to predict mortality and need for ICU admission in patients who present to the ED with suspected COVID-19 and may help patients and doctors in decision making.

BMJ

BACKGROUND

The COVID-19 pandemic is putting extraordinary pressure on emergency departments (EDs), clinical wards and intensive care units (ICUs). Clinical prediction models for COVID-19 outcomes have the potential to support decision making about hospital admission. Existing models that predict mortality for non-trauma patients presenting to the ED are unlikely to be well calibrated and optimally discriminating for patients with COVID-19.[1] Most currently available models specifically developed for patients with COVID-19 that were assessed with the prediction model risk of bias assessment tool contain a high risk of bias.[2–4] The most common reasons were non-representative selection of control patients, exclusion of patients in whom the event of interest was not observed by the end of the study, high risk of model overfitting and vague reporting. Additionally, the description of the study population or intended use of the models was often missing, and calibration of the model predictions was rarely assessed.

The recently proposed 4C Mortality Score is probably at low risk of bias, but was derived from a selected population of patients admitted to UK hospitals who were seriously ill (mortality rate of 32.2%). Predictors included the number of comorbidities and the Glasgow Coma Scale, items that are not easily and unambiguously obtained for patients with suspected COVID-19 at EDs everywhere.[5 6] Similarly, the promising risk scores Veterans Health Administration COVID-19 (VACO) and COVID-GRAM—predicting 30-day mortality in positively tested patients and critical illness in hospitalised patients, respectively—require knowledge on pre-existing comorbidities.[7 8] The COVID-GRAM model also requires chest radiography results.

We aimed to develop and validate a simple and valid model for predicting mortality and the need for ICU in all patients who are suspected to have COVID-19 when presenting at the ED. To facilitate implementation in clinical practice, we only included quickly and objectively obtainable patient characteristics, vital parameters and blood test values.

METHODS

Population

Nineteen large Dutch hospitals were requested to supply anonymised retrospective data on the cohorts of patients with COVID-19 who were admitted to their hospital. Of those hospitals, Catharina Hospital Eindhoven, Zuyderland Medical Center Heerlen, Isala Clinics Zwolle, Erasmus University Medical Center Rotterdam and Antonius Hospital Sneek supplied these data. The data from Antonius Hospital Sneek were not used in the analyses, because of large proportions of missing predictor values.

For model development, we used the data of patients who presented at the ED and were admitted to the hospital with suspected COVID-19 in the first wave of the pandemic, that is, from March up to and including August 2020. Patients being transferred to other hospitals were excluded since information on outcomes was missing. For model validation, we used data of patients who presented at the ED and were admitted to the hospital with suspected COVID-19 in the second wave of the pandemic, that is, from September up to and including December 2020. Potential multiple hospital admissions of the same patient were considered as independent hospital admissions.

Outcomes

The outcomes of interest were: (1) in-hospital death or transfer to a hospice within 28 days after hospital admission and (2) admission to ICU within 28 days after hospital admission.

Predictors

Based on prior literature, we included patient characteristics (sex, age, body mass index), vital parameters (oxygen saturation, systolic blood pressure, heart rate (HR), respiratory rate (RR), body temperature) and blood test values (C reactive protein (CRP), lactic dehydrogenase (LDH), D-Dimer, leucocytes, lymphocytes, monocytes, neutrophils, eosinophils, Mean Corpuscular Volume (MCV), albumin, bicarbonate, sodium, creatinine, urea), all measured at ED admission, as potential predictors.[2] Furthermore, we included the month of admission to capture potential changes in outcomes over time. In case of multiple measurements for the same patient, we used the first measurement after presentation at the ED. We used multivariate imputation by chained equations (R-packages mice) for multiple imputation of missing predictor values.[9 10] Multiple imputation in the validation data was undertaken separately from multiple imputation in the development data to ensure fully independent model validation.

Model development

Logistic regression was used to analyse associations between predictors and outcomes. We decided on including non-linear transformations of potential predictors on the basis of a full model with a restricted cubic spline (three knots; two regression coefficients) for each continuous predictor.[11 12] Based on Wald statistics, we selected the most promising predictors into a parsimonious model for easy use in clinical practice. To prevent overfitting, we used bootstrap validation—including the same variable selection strategy to mimic our modelling strategy—to estimate a uniform shrinkage factor.[12] The regression coefficients of the final model were multiplied by this shrinkage factor, and the model intercept was adjusted to ensure overall calibration of the model. We used the R-package rms (Regression Modelling Strategies) for regression analyses.[9 13]

Model validation

Model performance was assessed with temporal validation in second wave patients, in each of the four separate hospitals. We assessed discriminative ability with the area under the receiver operating characteristic curve (AUC)

and calibration with calibration plots of five equally sized groups of predicted risk, calibration intercepts and calibration slopes. The model-based concordance (mbc) was used to understand the impact of potential differences in casemix heterogeneity between the development and validation data on discriminative ability.[14]

Patient and public involvement

Patients were not directly involved in the design of this study. The outcome of interest and the potential predictors were selected up front by a group of hospital physicians caring for patients with COVID-19 (ED, internal medicine, pulmonary medicine, ICU). Since we retrospectively collected data, patients were not burdened by our study. In future research, we will convene multistakeholder panels of approximately 12 members including patients with COVID-19, relatives, hospitals physicians caring for patients with COVID-19, palliative care physicians and ethicists, with the aim to develop a full understanding of how the models may best support patients and clinicians in making critical patient-centred decisions.

RESULTS
Population and outcomes

The database contained 5912 patients who presented at the ED from March up to and including August 2020 and who were admitted to the hospital with a suspicion of COVID-19. Of those patients 81 (1.4%) were excluded because of a transfer to other hospitals (outcome not recorded). The development data included 5831 patients of whom 629 (10.8%) died, 5070 (86.9%) were discharged within 28 days after hospital admission, and 132 (2.3%) were still in hospital at 28 days after admission. Patients who died—in comparison with patients who were discharged—tended to be more often male (64% vs 56%), at older age (median 78 vs 69), with higher RR (median 23 vs 19) and HR (median 93 vs 90), lower oxygen saturation (median 94.1 vs 96.0), higher blood levels of CRP (median 91 vs 43), LDH (median 338 vs 237), creatinine (median 102 vs 82) and urea (median 9.6 vs 6.2) and lower blood levels of lymphocytes (median 0.80 vs 1.10) and albumin (median 36 vs 40) (table 1). Similar patterns were seen in 3252 patients who were admitted to hospital in the second wave of the pandemic from September up to and including December 2020, of whom 326 (10.0%) died, 2854 (87.8%) were discharged within 28 days after admission, and 72 (2.2%) were still in hospital at 28 days after admission.

Admission to ICU was fully recorded—including ICU admissions at a later time point than the initial hospital admission—for 2633 patients in 2 hospitals (214 ICU admissions within 28 days (8.1%)) in the first wave of the pandemic. Patients who were admitted to the ICU—in comparison with patients who were discharged or died without being admitted to the ICU—tended to be more often male (68% vs 57%), with higher RR (median 23 vs 19) and HR (median 91 vs 88), lower oxygen saturation (median 95.0 vs 95.8), higher blood levels of CRP (median 88 vs 47), LDH (median 318 vs 234), creatinine (median 93 vs 84) and urea (median 7.1 vs 6.6) and lower blood levels of albumin (median 38 vs 40) (table 2). In contrast with patients who died, patients who were admitted to the ICU were not older than patients who were discharged (median 68 vs 71), probably due to decisions not to admit frail patients to the ICU. Patterns were similar in 1466 patients (86 ICU admissions within 28 days (5.9%)) who were admitted to these 2 hospitals in the second wave of the pandemic.

Prediction of death

Patients who were admitted in the first month of the pandemic in the Netherlands, that is, in March 2020, were at substantially increased risk of death (table 3: multivariable OR 1.99; 95% CI 1.61 to 2.47). All models included this correction factor for the first month, to avoid overestimation of risk after the first month of the pandemic. Consequently, to avoid overestimation of the discriminative ability, we limited validation of models in the development data to patients who were admitted from April 2020 onward.

D-dimer concentration in the blood, measured to detect thrombosis, was not analysed in the regression analysis, because 64% and 76% were missing in the development and validation data, respectively (table 1). Based on a full model with restricted cubic splines of all potential variables, we decided to transform all biomarkers and RR with the natural logarithm, while keeping all other predictor effects linear. Some strong univariable associations with death—for example of logarithmically transformed lymphocytes and creatinine (table 3; Wald statistics 48 and 133, respectively)—were very weak in multivariable analysis (table 3; Wald statistics 0 and 4, respectively). The predictive ability of the resulting full multivariable regression model was mainly driven by age, LDH, urea, RR, CRP, Albumin, oxygen saturation and bicarbonate (ORs and Wald statistics in table 3). A simple model—named COVID outcome prediction in the emergency department (COPE)—with linear age and logarithmic transforms of RR, CRP, LDH, albumin and urea captured most of the ability to predict death within 28 days (table 3; figure 1). Based on internal bootstrap validation, we applied a shrinkage factor of 0.93 to the regression coefficients.

COPE showed good discrimination for predicting death in 4498 patients who were admitted from April up to and including August 2020 in the first wave (online supplemental figure 1); AUC in 4 hospitals 0.85 (95% CI 0.81 to 0.88)); 0.81 (95% CI 0.71 to 0.91); 0.86 (95% CI 0.82 to 0.90); 0.85 (95% CI 0.81 to 0.88)) and, more importantly, in the validation sample of 3235 patients who were admitted in the second wave from September up to and including December 2020 (figure 2; AUC in four hospitals: 0.82 (95% CI 0.78 to 0.86); 0.82 (95% CI 0.74 to 0.90); 0.79 (95% CI 0.70 to 0.88); 0.83 (95% CI 0.79 to 0.86)). The decrease in AUC over time was partly driven by less

Table 1 Baseline characteristics of development and validation patient cohorts median ('M') and quartile range ('Q1'=first quartile; 'Q3'=third quartile) are presented for all continuous variables

	% missing	All patients n=5831			Discharged n=5070			In hospital n=132			Dead n=629		
A. Development data													
Male sex	0	57 %			56 %			64 %			64 %		
		M	Q1	Q3	M	Q1	Q3	M	Q1	Q3	M	Q1	Q3
Age (years)	0	70	58	80	69	56	78	71	62	79	78	70	84
BMI (kg/m²)	58	26	23	30	26	23	30	25	23	29	26	23	30
HR (bpm)	39	90	78	103	90	78	103	87	76	99	93	80	107
SBP (mmHg)	42	133	118	150	133	119	151	136	115	152	131	114	145
RR (/min)	42	19	16	23	19	16	23	20	17	24	23	19	28
Saturation (%)	41	95.8	94.0	97.5	96.0	94.3	97.8	95.4	93.6	97.0	94.1	91.9	96.0
Temperature (°C)	40	37.3	36.7	38.1	37.3	36.7	38.0	37.5	37.0	38.1	37.4	36.7	38.1
CRP (mg/L)	7	48	10	118	43	8	110	85	34	160	91	41	180
D dimer (µg/L)	64	1100	527	2545	1028	504	2300	1950	719	8110	2100	949	4772
LDH (U/L)	18	244	200	322	237	197	302	300	234	422	338	253	492
Leucocytes (×10⁹/L)	7	9.1	6.7	12.7	9.1	6.7	12.6	9.4	6.5	12.5	9.2	6.4	14.0
Lymphocytes (×10⁹/L)	16	1.04	0.66	1.6	1.10	0.7	1.7	0.80	0.54	1.3	0.80	0.51	1.3
Albumin (g/L)	15	39	36	42	40	36	43	37	34	41	36	33	39
Bicarbonate (mmol/L)	45	23.6	21	26	24	22	26	23	21	26	22	20	25
Creatinine (µmol/L)	8	84	66	111	82	65	107	89	66	111	102	75	153
Eosinophils (×10⁹/L)	26	0.03	0.00	0.10	0.03	0.01	0.10	0.03	0.00	0.10	0.01	0.00	0.10
MCV (fL)	7	90	87	94	90	86	94	90	88	94	91	87	96
Monocytes (×10⁹/L)	30	0.67	0.44	0.95	0.68	0.45	0.95	0.59	0.33	0.83	0.61	0.33	0.94
Neutrophils (×10⁹/L)	16	5.6	2.2	9.0	5.6	2.2	8.9	7.0	4.0	10.3	5.5	1.8	9.1
Sodium (mmol/L)	9	138	135	140	138	135	140	137	133	141	137	134	140
Urea (mmol/L)	9	6.5	4.6	9.7	6.2	4.5	9.0	7.4	5.1	11.0	9.6	6.6	15.0
B. Validation data		n=3252			n=2854			n=72			n=326		
Male sex	0	56 %			56 %			46 %			61 %		
		M	Q1	Q3	M	Q1	Q3	M	Q1	Q3	M	Q1	Q3
Age (years)	0	71	58	80	69	55	79	72	58	82	79	73	85
BMI (kg/m²)	59	26	23	30	26	23	30	26	24	30	25	22	29
HR (bpm)	40	90	78	105	90	78	104	84	75	106	92	78	105
SBP (mmHg)	43	134	119	151	135	120	152	134	122	149	129	110	141

Continued

Table 1 Continued

	% missing	All patients			Discharged			In hospital			Dead		
RR (/min)	43	20	16	24	20	16	24	20	17	26	23	19	27
Saturation (%)	40	95.7	94.0	97.5	96.0	94.0	97.7	95.5	94.0	97.0	94.8	92.3	96.5
Temperature (°C)	42	37.3	36.7	38.1	37.3	36.7	38.1	37.3	36.8	38.1	37.2	36.4	38.0
CRP (mg/L)	9	57	16	124	54	15	120	76	21	169	80	33	159
D dimer (µg/L)	76	1060	531	2170	1013	490	2012	1080	640	2570	1495	870	3724
LDH (U/L)	22	247	203	334	242	200	317	281	226	390	315	238	489
Leucocytes ($\times 10^9$/L)	10	9.4	6.6	12.9	9.4	6.6	12.9	9.6	7.0	13.2	9.5	6.8	13.3
Lymphocytes ($\times 10^9$/L)	20	0.98	0.62	1.50	1.00	0.64	1.50	1.10	0.72	1.53	0.71	0.48	1.20
Albumin (g/L)	20	39	35	42	39	35	42	36	32	39	36	32	40
Bicarbonate (mmol/L)	50	23.5	21	26	23.6	21	26	22.8	20	27	22.8	20	25
Creatinine (µmol/L)	10	84	66	116	83	65	111	80	58	118	103	74	158
Eosinophils ($\times 10^9$/L)	27	0.03	0.01	0.10	0.03	0.01	0.10	0.03	0.01	0.09	0.03	0.00	0.04
MCV (fL)	10	90	87	94	90	87	94	92	88	95	91	88	96
Monocytes ($\times 10^9$/L)	30	0.67	0.43	0.98	0.67	0.44	0.98	0.57	0.38	1.00	0.58	0.36	0.97
Neutrophils ($\times 10^9$/L)	21	5.8	2.4	9.4	5.8	2.4	9.4	6.6	3.3	9.1	5.6	2.2	9.4
Sodium (mmol/L)	11	137	134	139	137	134	139	136	133	139	137	133	140
Urea (mmol/L)	11	6.9	4.9	10.4	6.6	4.7	9.8	7.3	5.3	11.8	10.3	7.4	17.0

Status, 'all patients', 'discharged', 'in hospital' and 'dead' is measured at 28 days after hospital admission.
BMI, body mass index; CRP, C reactive protein; HR, heart rate; LDH, lactic dehydrogenase; RR, respiratory rate; SBP, systolic blood pressure.

Table 2 Baseline characteristics of development and validation patient cohorts of two hospitals with a full record of ICU admissions median ('M') and quartile range ('Q1'=first quartile; 'Q3'=third quartile) are presented for all continuous variables

	% missing	All patients n=2633			Discharged or dead n=2387			In hospital n=32			ICU admission n=214		
A. Development data													
Male sex	0	58 %			57 %			63 %			68 %		
		M	Q1	Q3	M	Q1	Q3	M	Q1	Q3	M	Q1	Q3
Age (years)	0	71	58	80	71	57	80	80	71	85	68	59	74
BMI (kg/m²)	48	26	23	30	26	23	30	25	22	29	27	24	31
HR (bpm)	14	88	77	100	88	77	100	87	77	98	91	79	104
SBP (mmHg)	20	131	116	149	131	116	149	140	122	159	129	114	145
RR (/min)	15	19	15	23	19	15	23	18	16	22	23	18	27
Saturation (%)	10	95.7	94.0	97.3	95.8	94.0	97.4	96.0	95.0	97.5	95.0	93.1	96.7
Temperature (°C)	11	37.2	36.7	37.9	37.2	36.7	37.9	37.4	36.7	37.8	37.4	36.7	38.2
CRP (mg/L)	4	49	11	120	47	10	110	82	41	175	88	20	178
D dimer (µg/L)	79	1100	463	2700	1040	460	2453	3400	2000	5400	2000	590	4900
LDH (U/L)	11	239	197	317	234	194	306	270	216	324	318	234	444
Leucocytes (×10⁹/L)	4	9.5	6.7	13.2	9.4	6.8	13.1	12.0	9.2	16.1	10.3	6.3	14.7
Lymphocytes (×10⁹/L)	11	1.00	0.67	1.6	1.00	0.7	1.6	0.74	0.55	1.3	1.10	0.7	1.5
Albumin (g/L)	8	40	36	43	40	36	43	37	34	41	38	35	42
Bicarbonate (mmol/L)	66	23	21	25	23	21	25	24	22	26	21	19	24
Creatinine (µmol/L)	5	85	67	113	84	67	112	85	64	113	93	73	125
Eosinophils (×10⁹/L)	27	0.1	0.03	0.10	0.1	0.03	0.11	0.03	0.03	0.10	0.1	0.03	0.10
MCV (fL)	5	91	87	95	91	87	95	92	90	97	91	88	95
Monocytes (×10⁹/L)	41	0.72	0.47	1.00	0.73	0.48	1.00	0.71	0.56	1.02	0.64	0.39	1.00
Neutrophils (×10⁹/L)	11	7.2	4.8	10.5	7.1	4.8	10.4	9.8	6.6	14.5	7.7	5.0	11.1
Sodium (mmol/L)	5	137	134	140	137	134	140	138	134	140	137	134	139
Urea (mmol/L)	5	6.6	4.7	9.9	6.6	4.6	9.8	8.4	6.3	11.9	7.1	5.1	11.0
B. Validation data		n=1466			n=1356			n=24			n=86		
Male sex	0	58 %			58 %			38 %			71 %		
		M	Q1	Q3	M	Q1	Q3	M	Q1	Q3	M	Q1	Q3
Age (years)	0	71	57	80	71	56	80	80	71	80	70	62	76
BMI (kg/m²)	40	26	23	30	26	23	30	25	22	31	28	24	30
HR (bpm)	16	89	77	103	89	77	103	78	73	99	89	78	103
SBP (mm Hg)	21	134	119	151	134	119	151	132	111	149	141	124	157

Continued

Table 2 Continued

	% missing	All patients			Discharged or dead			In hospital			ICU admission		
RR (/min)	17	19	16	23	19	16	23	19	16	22	23	19	28
Saturation (%)	9	95.7	94.0	97.3	95.9	94.0	97.5	96.2	94.5	97.0	93.8	91.3	95.7
Temperature (°C)	13	37.2	36.7	37.9	37.2	36.7	37.9	37.2	36.7	37.7	37.4	37.0	38.0
CRP (mg/L)	5	49	14	115	47	13	110	77	25	135	114	40	206
D dimer (µg/L)	95	460	210	1275	440	210	1350	770	770	770	913	890	937
LDH (U/L)	13	238	196	314	234	195	307	254	217	288	360	234	531
Leucocytes (×10^9/L)	6	9.7	6.8	13.2	9.7	6.8	13.2	12.0	7.3	15.0	8.9	6.7	13.1
Lymphocytes (×10^9/L)	12	1.00	0.65	1.50	1.00	0.66	1.50	1.10	0.80	1.33	0.84	0.60	1.55
Albumin (g/L)	9	40	37	42	40	37	42	37	35	38	39	36	42
Bicarbonate (mmol/L)	65	23	21	25	23	21	25	23	21	26	22	19	25
Creatinine (µmol/L)	6	86	68	119	85	68	117	80	58	122	96	77	138
Eosinophils (×10^9/L)	20	0.07	0.03	0.10	0.07	0.03	0.11	0.05	0.03	0.10	0.05	0.03	0.10
MCV (fL)	6	91	88	95	91	88	95	94	90	97	91	87	95
Monocytes (×10^9/L)	33	0.70	0.47	1.00	0.70	0.48	1.00	0.68	0.48	1.03	0.61	0.40	1.00
Neutrophils (×10^9/L)	13	7.4	4.8	11.0	7.3	4.7	11.0	9.0	5.5	13.1	7.4	5.0	10.8
Sodium (mmol/L)	6	137	134	139	137	134	139	136	133	138	135	132	138
Urea (mmol/L)	7	6.7	4.9	10.0	6.6	4.8	9.9	8.0	4.7	12.8	7.8	5.9	12.2

Status, 'all patients', 'discharged or dead' (without being admitted to the ICU), 'in hospital' and 'ICU admission' is measured at 28 days after hospital admission.
BMI, body mass index; CRP, C reactive protein; HR, heart rate; ICU, intensive care unit; LDH, lactic dehydrogenase; RR, respiratory rate; SBP, systolic blood pressure.

Table 3 Univariable and multivariable associations between predictors and death within 28 days OR with 95% CIs for separate variables (columns 'univariable'), for a model with all available predictors (columns 'full model') and for a model with only the six strongest predictors (columns 'selected model')

Predictor	Contrast	Univariable			Full model			Selected model			COPE
		OR	95% CI	Wald	OR	95% CI	Wald	OR	95% CI	Wald	OR
Month	≥April vs <April	2.57	2.16 to 3.05	114	1.99	1.61 to 2.47	39	2.06	1.68 to 2.52	49	1.96
Sex	Male vs female	1.37	1.15 to 1.63	13	1.12	0.9 to 1.39	1				
Age (years)	80 vs 58	3.07	2.63 to 3.58	201	3.16	2.56 to 3.91	113	2.95	2.42 to 3.6	112	2.74
BMI (kg/m²)	35 vs 25	1.07	0.89 to 1.28	1	1.09	0.9 to 1.34	1				
HR (bpm)	103 vs 78	1.16	1.05 to 1.29	8	1.19	0.97 to 1.45	3				
SBP (mm Hg)	150 vs 118	0.76	0.65 to 0.89	12	0.86	0.73 to 1.01	3				
RR (/min)	23 vs 16	1.98	1.77 to 2.21	150	1.63	1.34 to 1.99	24	1.91	1.63 to 2.23	64	1.82
Saturation (%)	97.5 vs 94	0.61	0.52 to 0.72	37	0.77	0.65 to 0.9	10				
Temperature (°C)	38 vs 37	1.16	1.02 to 1.32	5	1.05	0.87 to 1.27	0				
CRP (mg/L)	118 vs 10	2.76	2.35 to 3.25	149	1.54	1.22 to 1.93	14	1.57	1.27 to 1.93	18	1.52
LDH (U/L)	322 vs 200	2.17	1.99 to 2.36	309	1.83	1.62 to 2.06	99	1.85	1.66 to 2.05	127	1.77
Leucocytes (×10⁹/L)	12.7 vs 6.7	1.01	0.92 to 1.11	0	0.88	0.75 to 1.02	3				
Lymphocytes (×10⁹/L)	1.6 vs 0.66	0.67	0.6 to 0.75	48	1.03	0.9 to 1.19	0				
Albumin (g/L)	42 vs 36	0.58	0.53 to 0.62	191	0.8	0.71 to 0.9	14	0.77	0.69 to 0.86	21	0.78
Bicarbonate (mmol/L)	25.9 vs 21.4	0.71	0.65 to 0.78	54	0.81	0.71 to 0.92	10				
Creatinine (µmol/L)	111 vs 66	1.58	1.46 to 1.71	133	0.85	0.72 to 1	4				
Eosinophils (×10⁹/L)	0.1 vs 0.004	0.77	0.64 to 0.93	7	1.08	0.9 to 1.3	1				
MCV (fL)	94 vs 87	1.18	1.08 to 1.29	14	1.1	0.99 to 1.23	3				
Monocytes (×10⁹/L)	0.95 vs 0.44	0.84	0.74 to 0.96	7	1.08	0.87 to 1.35	1				
Neutrophils (×10⁹/L)	9 vs 2.2	0.97	0.88 to 1.07	0	0.94	0.82 to 1.08	1				
Sodium (mmol/L)	140 vs 135	0.98	0.96 to 1.01	1	1.03	0.98 to 1.09	1				
Urea (mmol/L)	9.7 vs 4.6	2.48	2.24 to 2.76	291	1.79	1.43 to 2.24	26	1.61	1.42 to 1.83	53	1.56

Variable importance is expressed with the Wald statistic (columns 'Wald'). The ORs of the final model are based on the selected model, with a uniform shrinkage factor of 0.93 (column 'COPE'). Associations are based on 5831 patients of whom 629 died within 28 days.
BMI, body mass index; COPE, COVID-19 outcome prediction in the emergency department; CRP, C reactive protein; HR, heart rate; LDH, lactic dehydrogenase; RR, respiratory rate; SBP, systolic blood pressure.
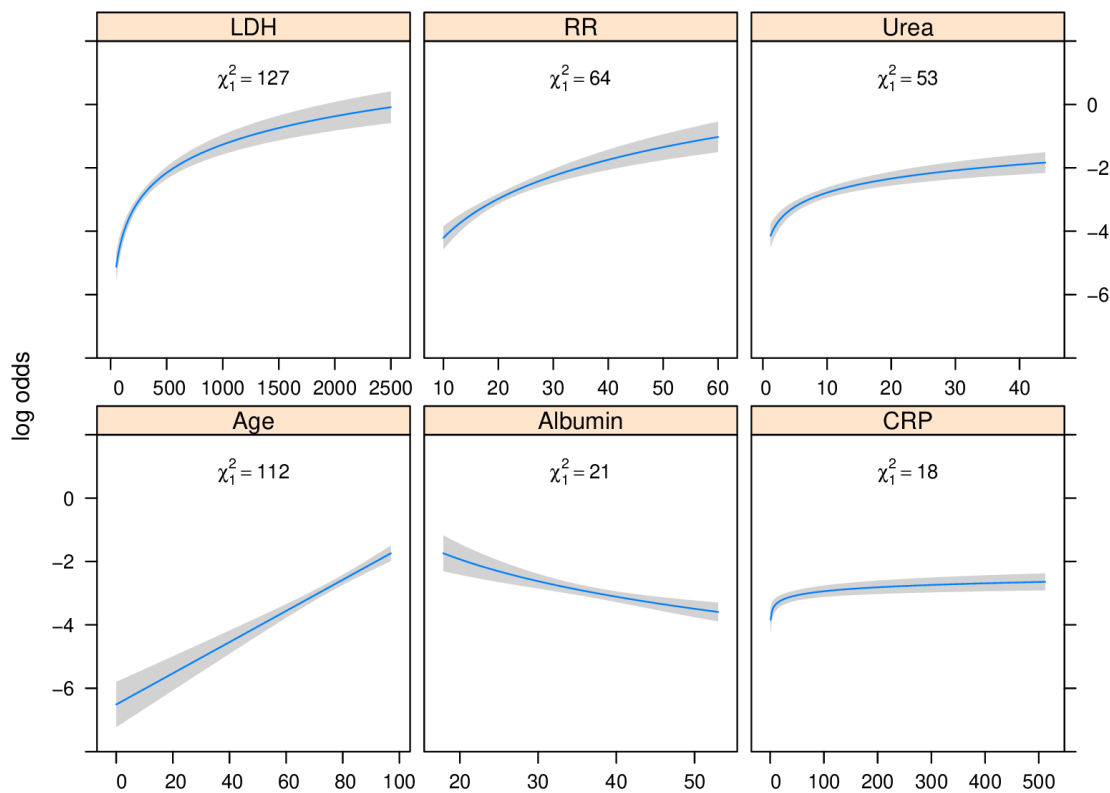
Figure 1 Multivariable effects of continuous predictors of death within 28 days predictions of the logarithm of the odds by continuous predictor levels, with other predictor levels set to the median. Wald statistics are listed within each plot to express variable importance (higher is better). CRP, C reactive protein; LDH, lactic dehydrogenase; RR, respiratory rate.

case mix heterogeneity—expressed by a lower model-based AUC (mbc)—of second wave patients (figure 2; mbc in four hospitals: 0.81; 0.82, 0.81; 0.82) as compared with first wave patients (online supplemental figure 1); mbc in four hospitals 0.82; 0.85, 0.83, 0.84). COPE was well calibrated in second wave patients of each of the four hospitals, both on average—expressed by hospital-specific calibration intercepts: 0.08 (95% CI −0.15 to 0.30); −0.17 (95% CI −0.65 to 0.30); −0.01 (95% CI −0.40 to 0.39); −0.12 (95% CI −0.30 to 0.07)—and by predicted risk levels—expressed by hospital-specific calibration slopes: 1.09 (95% CI 0.86 to 1.31); 0.90 (95% CI 0.49 to 1.32); 0.91 (95% CI 0.57 to 1.25); 0.97 (95% CI 0.79 to 1.14) (figure 2).

When stratifying second wave patients according to a mortality risk threshold equal to the event rate (10%), COPE assigned high risk to 246/326 patients who actually died (76% sensitivity, ie, 24% false negatives) and low risk to 2086/2926 patients who actually survived (71% specificity, ie, 29% false positives). With a 5% risk threshold, the sensitivity increased to 93% while the specificity decreased to 49%. Based on a 20% risk threshold, the sensitivity decreased to 49% while the specificity increased to 89%.

Prediction of need for ICU admission

The probability of being admitted to the ICU was decreasing with age after the age of 70 (OR of age ≥80 vs 70–79: 0.17 (95% CI 0.10 to 0.30)), likely reflecting the decision not to admit older patients to the ICU. When adjusting for this decreasing age effect after the age of 70—by including a

linear spline with a knot at age 70 in the regression model (online supplemental figure 2)—the strongest predictors of death were also predictive of ICU admission within 28 days, but associations were generally weaker for the latter (table 4 vs table 3). In patients below the age of 70, admitted from April up to and including August 2020, a model with the linear predictor of death calibrated to ICU admission had similar discriminative ability to a model that refitted all the predictor effects (AUC 0.71 for both models). For robustness, we implemented the calibrated model, also adjusted for a linearly decreasing age effect after the age of 70, and not the refitted model (calibration slope 0.60; 95% CI 0.49 to 0.70) into COPE for predicting ICU admission. To predict the need for ICU admission of future patients over the age of 70 COPE ignores the decreasing age effect after the age of 70, since the observed ICU admission rate is probably an inaccurate estimate of the medical need for ICU admission. By fitting a linearly decreasing age effect in patients over the age of 70 which is not applied when predicting for future patients, predictions of ICU admission after the age of 70 are based on an extrapolation of the observed age effect on ICU admission in patients below the age of 70. Due to the weaker predictor effects, the discriminative ability of COPE was more moderate for predicting ICU admission than for predicting death (online supplemental figure 3); AUC in two hospitals: 0.66 (95% CI 0.58 to 0.74); 0.79 (95% CI 0.69 to 0.88)). Although COPE significantly overestimated ICU admission in second wave patients (figure 3; calibration
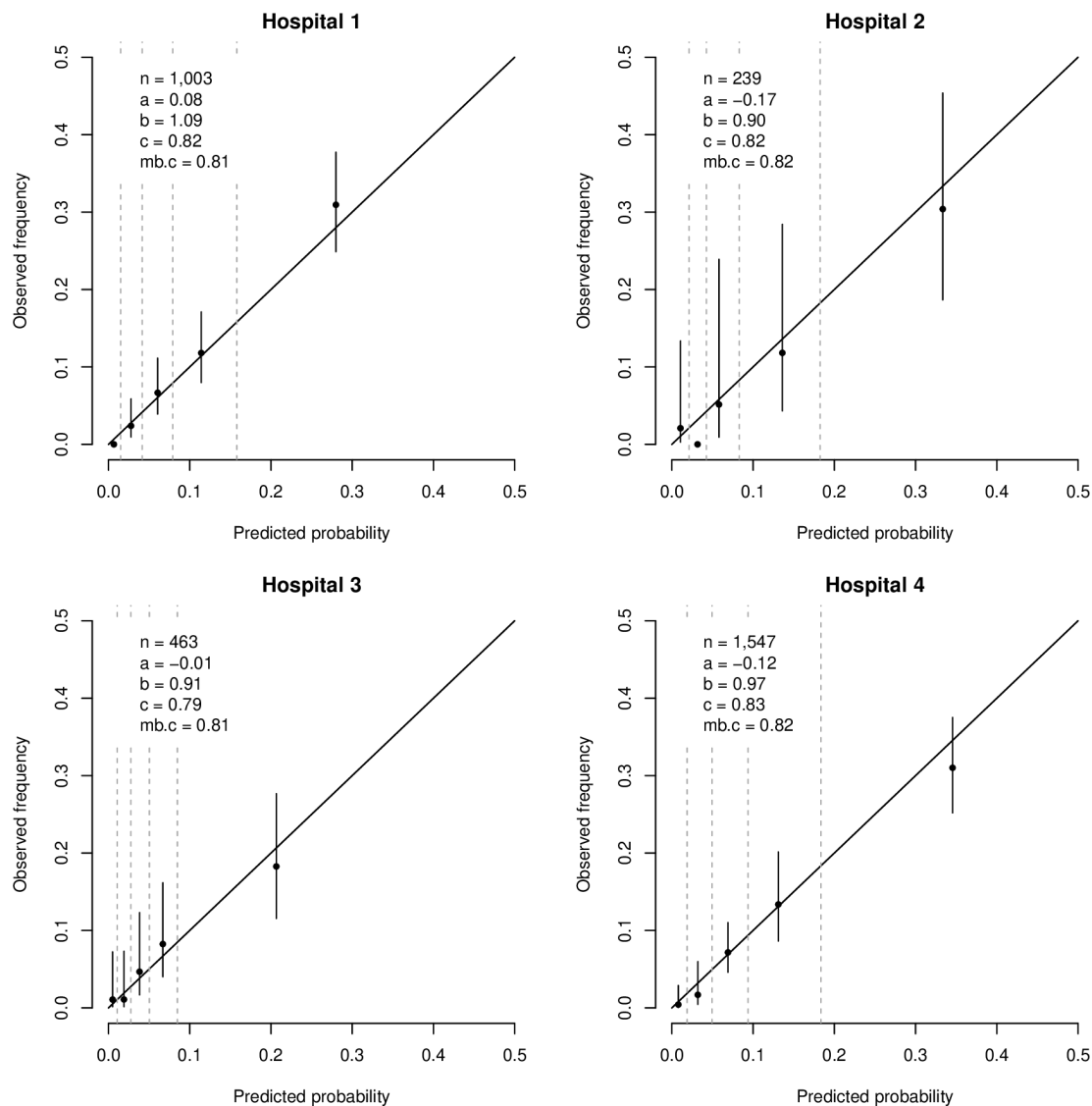
Figure 2 Temporal validation: Performance of COPE for predicting death in second wave patients calibration plots of patients who were admitted since September 2020 in four separate Dutch hospitals. n is number of patients; a=calibration intercept (0 is perfect); b=calibration slope (1 is perfect); c=AUC (0.5 is useless; 1 is perfect); mbc=model-based AUC. AUC, area under the curve; COPE, COVID-19 outcome prediction in the emergency department.

intercept in two hospitals: −0.41 (95% CI −0.77 to −0.05); −0.72 (95% CI −1.34 to −0.11)), it was better able to identify the patients at high risk of needing ICU admission, as expressed by higher discriminative ability (figure 3; AUC in two hospitals: 0.84 (95% CI 0.78 to 0.90); 0.81 (95% CI 0.66 to 0.95)) and substantially stronger predictor effects (calibration slope in two hospitals: 1.55 (95% CI 1.03 to 2.06); 1.53 (95% CI 0.60 to 2.46)).

Model presentation
The resulting COPE models for predicting death as well as need for ICU admission within 28 days after hospital admission (formulas in table 5) are implemented as a publicly accessible web-based application (https://mdmerasmusmc. shinyapps.io/COPE/) and as independent mobile apps ('COPE Decision Support'). For optimal transparency, the web and mobile applications include a detailed description of the derivation of COPE (online supplemental file 1),

descriptions of the data that were used for development and validation of COPE, and calibration plots of temporal validation in the separate hospitals.

According to the Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis checklist (online supplemental table 1), all relevant items are covered in this manuscript, except for the availability of data sets.[15 16] The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to data transfer agreements with each of the contributing hospitals.

DISCUSSION
We developed COPE for prediction of in-hospital death and need for intensive care when patients with suspected COVID-19 present at the ED. Developed using patient

Table 4 Multivariable associations between predictors and ICU admission within 28 days OR with 95% CIs for separate variables (columns 'univariable') and for a model with the six strongest predictors of death, corrected for a decreasing probability of ICU admission after the age of 70 (columns 'multivariable')

Predictor	Contrast	Univariable			Multivariable		
		OR	95% CI	Wald	OR	95% CI	Wald
Month	≥April vs <April	2.06	1.51 to 2.81	21	1.63	1.16 to 2.28	8
Age (years)	80 vs 58	1.96	1.47 to 2.62	21	1.76	1.32 to 2.35	15
RR (/min)	23 vs 16	1.76	1.48 to 2.09	40	1.71	1.4 to 2.09	27
CRP (mg/L)	118 vs 10	1.88	1.44 to 2.44	22	1.3	0.95 to 1.77	3
LDH (U/L)	322 vs 200	1.73	1.52 to 1.98	66	1.44	1.25 to 1.67	24
Albumin (g/L)	42 vs 36	0.75	0.64 to 0.88	13	0.95	0.78 to 1.17	0
Urea (mmol/L)	9.7 vs 4.6	1.29	1.08 to 1.54	8	1.36	1.1 to 1.66	8
Adjusted for							
Max (age-70, 0) (years)*	80 vs 58	0.2	0.13 to 0.30	57	0.17	0.11 to 0.27	65

Associations are based on 2633 patients of whom 214 were admitted to the ICU within 28 days.
Variable importance is expressed with the Wald statistic (columns 'Wald').
*In univariable regression, only the association of age with ICU admission was adjusted for max (age-70, 0).
CRP, C reactive protein; ICU, intensive care unit; LDH, lactic dehydrogenase; RR, respiratory rate.

data from the first wave of the pandemic, based on six quickly and objectively obtainable predictors when entering the ED—age, RR, LDH, CRP, albumin and urea—COPE discriminated well and was well calibrated in patients admitted to hospitals in the second wave of the pandemic, both for predicting in-hospital death and for ICU admission.

The clinical presentation of COVID-19 is broad and varies from asymptomatic to critical disease. Some patients who initially have mild symptoms progress to severe disease within 1 week.[17] In the ED physicians need to identify high-risk patients—that is, those at high risk of deterioration and/or death—requiring treatment in the ICU, intermediate-risk patients requiring admission

to the clinical ward, and low-risk patients who can potentially be sent home. Since COPE is based on data that are routinely measured, or at least readily available in the ED, it can act as a tool to support such decisions. Hospitalised patients who are at high risk for mortality or need for ICU admission should be more intensively watched, and when a high load of high-risk patients occurs in the ED, this should be taken into account in the ICU capacity planning. COPE does not explicitly define treatment decisions based on risk thresholds, such as: send the patient home when the mortality risk is below a risk threshold of x%, or: admit the patient to the ICU over a mortality risk threshold of y%. These currently unavailable risk thresholds, and hence the resulting treatment decisions, depend on a
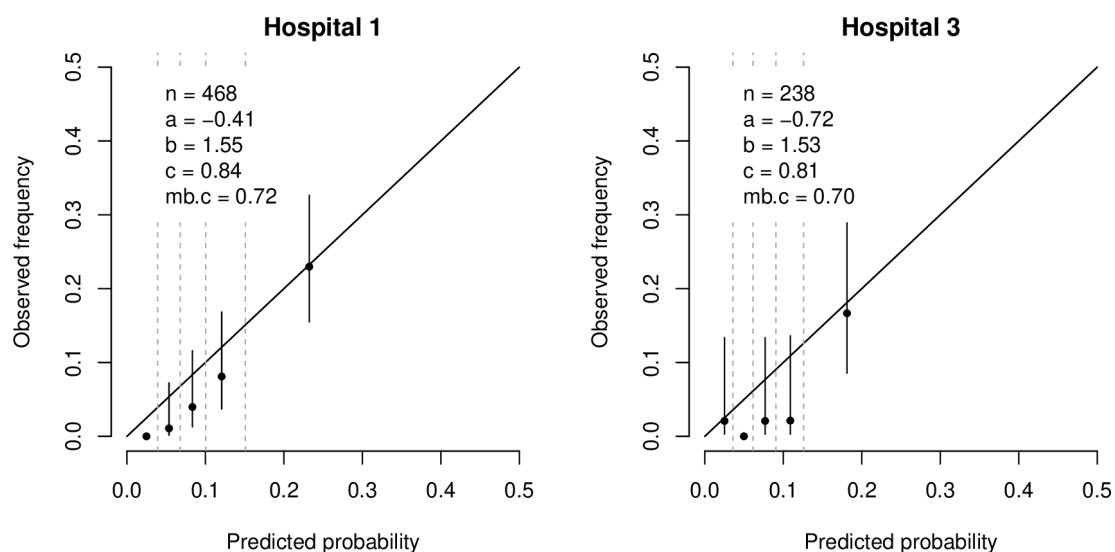
Figure 3 Temporal validation: performance of COPE for predicting ICU admission in second wave patients calibration plots of patients who were admitted since September 2020 in two separate Dutch hospitals. n is number of patients; a=calibration intercept (0 is perfect); b=calibration slope (1 is perfect); c=AUC (0.5 is useless; 1 is perfect); mbc=model-based AUC. AUC, area under the curve; COPE, COVID-19 outcome prediction in the emergency department; ICU, intensive care unit.

Table 5 COPE definition

Predictor	Minimum	Maximum
Age (years)	0	100
RR (/min)	10	60
CRP (mg/L)	1	500
LDH (U/L)	50	4000
Albumin (g/L)	10	60
Urea (mmol/L)	1	80

lp=−13.6 +

0.04575×age +
1.654×log (RR) +
0.1688×log (CRP) +
1.197×log (LDH) +
−1.585×log (albumin) +
0.5953×log (urea)

Probability of death within 28 days=1/(1+exp(-lp))

Probability of ICU admission within 28 days=1/(1+exp(−(−0.08949+0.5970 × lp)))

Implemented in web application https://mdmerasmusmc.shinyapps.io/COPE/ and mobile application 'COPE Decision Support'.
COPE, COVID-19 outcome prediction in the emergency department; CRP, C reactive protein; ICU, intensive care unit; LDH, lactic dehydrogenase; RR, respiratory rate.

trade-off between benefits and harms (including costs) of hospital or ICU admission.[18] Further research is necessary to better understand the benefits and harms of hospital admission and of ICU admission, for individual patients with COVID-19.[19] Furthermore, treatment decisions may depend on the availability of resources. The decision to admit a patient to the hospital, or even to the ICU, may depend on the availability of hospital beds and ICU beds. Likewise, the decision to send a low-risk patient home may depend on the availability of relatives who are willing to care for the patient at home. Although it is currently not possible to define explicit risk-based treatment decisions for patients with COVID-19, the risk predictions provided by COPE can be factored in by doctors, patients and relatives, when making decisions about hospital or ICU admission.

We requested 19 large Dutch hospitals to supply anonymised retrospective data on the cohorts of patients with COVID-19 who were admitted to their hospital. This request for data was sent out very early in the pandemic and was greeted with enthusiasm. Probably due to the enormous pressure on healthcare at that time, four hospitals supplied useable data for the analysis. The contributing hospitals were well spread over the Netherlands, with one in the west, two in the south and one in the east of the country and are a mix of academic and large teaching hospitals. we believe they are representative for healthcare in the Netherlands. Although the consistently good performance of COPE across the hospitals may support its generalisability to other countries, geographical validation would be additionally reassuring, since the

epidemic, and clinical practice—for example, access to ICU or other enhanced care—for this novel disease, may have substantial intercountry variability.

COPE was developed based on 5831 patients of whom 629 died within 28 days. This effective sample size of 629 events was ample to start the development process with a full model of 45 regression coefficients (14 events per variable), that is, one binary predictor (sex) and 22 continuous predictors with 2 regression coefficients—due to using non-linear terms—each.[20] To prevent too extreme predictions of COPE in new data, we applied a shrinkage factor to its regression coefficients, based on a bootstrap procedure with backward selection starting from the full model.[12]

Our explicit aim was to develop a score based on quickly and objectively obtainable predictors at presentation at the ED. Consequently, pre-existing comorbidities, the level of consciousness measured by the Glasgow Coma Scale, and chest radiography results—although predictive for outcomes of patients with COVID-19 in other studies—were not considered here.[5 7 8] Some predictors were promising in univariable analysis, such as lymphocytes and creatinine, but had negligible effects in multivariable analysis, because of strong correlations with other, more important predictors. Other predictors, such as oxygen saturation and bicarbonate, were significantly associated with death in multivariable analysis, but were not selected into the final model, since our explicit aim was to develop a simple model and the incremental value of these predictors was minimal. To achieve this aim, we only selected the strongest predictors—age, RR, LDH, CRP, albumin and urea—resulting in a parsimonious but well-performing model.

We aimed to predict outcomes for all patients who present to the ED with suspected COVID-19, regardless of actual hospital admission. Our data were limited to patients who presented at the ED and were admitted to hospital, because their outcomes were captured in the retrospective hospital database, while outcomes of patients who were sent home were not captured in the retrospective hospital database. Nevertheless, over 90% of the patients who presented to the ED with suspected COVID-19 were admitted to hospital and it is reasonable to assume that our predictions can be extrapolated to the less than 10% of patients who were sent home. Of note, the discriminative ability of our model is probably better in all patients presenting to the ED, due to a more heterogeneous casemix: patients who were sent home are likely to have more favourable predictor levels and more favourable outcomes than patients who are admitted.[14]

Besides mortality, we aimed to predict the need for ICU admission. A limitation of our study is that the need for ICU admission differs from the observed decisions on ICU admission, and is inherently difficult to model, because recorded ICU admissions express historical decisions at national, regional, hospital or even intensivist level. As a robust solution, we exploited the strong correlation between need for intensive care and death, by calibrating

our model for predicting death to the observed ICU admissions, adjusting for a linear decrease with age after the age of 70. Hence, we assumed a linear relationship between (the logarithm of the odds of) death and need for ICU admission, and that all patients below the age of 70 needing intensive care were actually admitted to the ICU, that is, the need for ICU admission is well estimated by the observed decisions on ICU admission for patients below the age of 70. The latter is reasonable given the sufficiency of ICU beds for Dutch patients throughout the pandemic. The discriminative ability of this recalibration approach was very similar to that of a model that refitted all associations between COPE predictors and ICU admission. With temporal validation in two separate hospitals, we showed that COPE discriminated very well between patients at low and high risk of ICU admission and that the predicted probability of ICU admission was well calibrated for the 20% highest-risk patients (highest-risk quintiles in figure 3). Nevertheless, recalibration of COPE for predicting need for ICU admission to local circumstances may be necessary.

The absence of external validation in our study—measuring the predictive performance of COPE in hospitals that were not present in the development data—may be considered a limitation of this study.[21] However, the combination of temporal validation—in second wave patients—and geographical validation—in separate hospitals—is a strength of this study.[22] Although COPE already performed very well when validated across time and space, future research should focus on analyses of potential time trends not captured by the predictors—for example, changes in mortality due to: improvements in treating patients with COVID-19; mutations of COVID-19; changes in patient casemix or critical care capacity fluctuations[23]—potential changes in predictor effects in time (interactions between predictors and time), and the impact of potential differences in patient case mix and differences in clinical care in countries other than the Netherlands (international validation). These casemix and clinical care differences should primarily affect calibration, requiring an update of the model intercept, but not discrimination. The emergence of new COVID-19 variants with potentially different mortality risk may especially require frequent analyses of the need for model updating.[24]

In conclusion, COPE, a simple tool based on six quickly and objectively obtainable predictors in the ED, is well able to predict mortality and need for ICU admission for patients who present to the ED with suspected COVID-19. COPE may support patients and doctors in decision making.

Author affiliations
[1]Department of Public Health, Erasmus MC, Rotterdam, The Netherlands
[2]Predictive Analytics and Comparative Effectiveness Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, Massachusetts, USA
[3]Department of Medical Informatics, Erasmus MC, Rotterdam, The Netherlands
[4]Department of Internal Medicine, Erasmus MC, Rotterdam, The Netherlands
[5]Emergency Department, Erasmus MC, Rotterdam, The Netherlands
[6]Department of Intensive Care, Catharina Hospital, Eindhoven, The Netherlands
[7]Department of Intensive Care, Zuyderland Medical Centre Heerlen, Heerlen, The Netherlands
[8]Department of Traumatology, Maastricht University Medical Centre+, Maastricht, The Netherlands
[9]Department of Internal Medicine, Antonius Hospital Sneek, Sneek, The Netherlands
[10]Department of Intensive Care, Antonius Hospital Sneek, Sneek, The Netherlands
[11]Department of Pulmonary Medicine, Isala Hospitals, Zwolle, The Netherlands
[12]Department of Intensive Care, Erasmus MC, Rotterdam, The Netherlands
[13]Department of Pulmonary Medicine, Erasmus MC, Rotterdam, The Netherlands
[14]Department of Medical Microbiology and Infectious Diseases, Erasmus MC, Rotterdam, The Netherlands
[15]Department of Internal Medicine, Department of Medical Microbiology and Infectious Diseases, Erasmus MC, Rotterdam, The Netherlands
[16]Executive Board, UMCG, Groningen, The Netherlands

ORCID iDs
David van Klaveren http://orcid.org/0000-0002-2096-606X
Rozemarijn L van Bruchem-Visser http://orcid.org/0000-0002-6833-7641

REFERENCES

1 Brink A, Alsma J, Fortuin AW, et al. Prediction models for mortality in adult patients visiting the emergency department: a systematic review. *Acute Med* 2019;18:171–83.
2 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.
3 Sperrin M, Grant SW, Peek N. Prediction models for diagnosis and prognosis in Covid-19. *BMJ* 2020;369:m1464.
4 Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med* 2019;170:W1.
5 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.
6 Docherty AB, Harrison EM, Green CA, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. *BMJ* 2020;369:m1985.
7 Liang W, Liang H, Ou L, et al. Development and validation of a clinical risk score to predict the occurrence of critical illness in hospitalized patients with COVID-19. *JAMA Intern Med* 2020;180:1081–9.
8 King JT, Yoon JS, Rentsch CT, et al. Development and validation of a 30-day mortality index based on pre-existing medical administrative data from 13,323 COVID-19 patients: the Veterans health administration COVID-19 (VACO) index. *PLoS One* 2020;15:e0241825.
9 R Core Team. R: a language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria, 2020. Available: http://www.R-project.org/
10 Buuren Svan, Groothuis-Oudshoorn K. mice : multivariate imputation by chained equations in R. *J Stat Softw* 2011;45:1–67.
11 Stone CJ, Koo C-Y. Additive splines in statistics. *Proceedings of the American Statistical Association Original pagination is p*, 1985.
12 Harrell FE. *Regression modeling strategies: with applications to linear models, logistic and ordinal regression and survival analysis*. Springer International Publishing, 2015.
13 Harrell FE. rms: regression modeling strategies. R package version 5.1-3.1, 2019. Available: http://CRAN.R-project.org/package=rms
14 van Klaveren D, Gönen M, Steyerberg EW, et al. A new concordance measure for risk prediction models in external validation settings. *Stat Med* 2016;35:4136–52.
15 Collins GS, Reitsma JB, Altman DG, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *Ann Intern Med* 2015;162:55–63.
16 Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–73.
17 Cohen PA, Hall LE, John JN, et al. The early natural history of SARS-CoV-2 infection: clinical observations from an urban, ambulatory COVID-19 clinic. *Mayo Clin Proc* 2020;95:1124–6.
18 Gold MR, Siegel JE, Russell LB, eds. *Cost-effectiveness in health and medicine New York*. NY: Oxford University Press, 1996.
19 Ioannidis JPA, Garber AM. Individualized cost-effectiveness analysis. *PLoS Med* 2011;8:e1001058.
20 Steyerberg EW, Harrell FE. Prediction models need appropriate internal, internal-external, and external validation. *J Clin Epidemiol* 2016;69:245–7.
21 Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med* 2000;19:453–73.
22 Austin PC, van Klaveren D, Vergouwe Y, et al. Geographic and temporal validity of prediction models: different approaches were useful to examine model performance. *J Clin Epidemiol* 2016;79:76–85.
23 RECOVERY Collaborative Group, Horby P, Lim WS, et al. Dexamethasone in hospitalized patients with Covid-19. *N Engl J Med* 2021;384:693–704.
24 Challen R, Brooks-Pollock E, Read JM, et al. Risk of mortality in patients infected with SARS-CoV-2 variant of concern 202012/1: matched cohort study. *BMJ* 2021;372:n579.