Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study

Steve Goodacre, Professor of Emergency Medicine, University of Sheffield, UK, s.goodacre@sheffield.ac.uk

Ben Thomas, Study Manager, University of Sheffield, UK, b.d.thomas@sheffield.ac.uk

Laura Sutton, Statistician/Research Associate, University of Sheffield, UK, l.j.sutton@sheffield.ac.uk

Matthew Bursnall, Medical Statistician, University of Sheffield, UK, m.bursnall@sheffield.ac.uk

Ellen Lee, Statistician, University of Sheffield, UK, e.lee@sheffield.ac.uk

Mike Bradburn, Senior Medical Statistician, m.bradburn@sheffield.ac.uk

Amanda Loban, Head of Data Management, University of Sheffield, UK, a.loban@sheffield.ac.uk

Simon Waterhouse, Lead Data Specialist, University of Sheffield, UK, s.waterhouse@sheffield.ac.uk

Richard Simmonds, Data Management/Information Systems Officer, University of Sheffield, UK, r.p.simmonds@sheffield.ac.uk

Katie Biggs, Assistant Director, University of Sheffield, UK, c.e.biggs@sheffield.ac.uk

Carl Marincowitz, Clinical Lecturer in Emergency Medicine , University of Sheffield, UK, c.marincowitz@sheffield.ac.uk

Jose Schutter, Research Assistant, University of Sheffield, UK, j.schutter@sheffield.ac.uk

Sarah Connelly, Research Assistant, University of Sheffield, UK, s.l.connelly@sheffield.ac.uk

Elena Sheldon, Research Assistant, University of Sheffield, UK, e.m.sheldon@sheffield.ac.uk

Jamie Hall, Research Assistant, University of Sheffield, UK, jamie.hall@sheffield.ac.uk

Emma Young, Research Assistant, University of Sheffield, UK, e.l.young@sheffield.ac.uk

Andrew Bentley, Consultant in ICM & Respiratory Medicine, Manchester University NHS Foundation Trust, Wythenshawe Hospital, UK, andrew.bentley@manchester.ac.uk

Kirsty Challen, Consultant in Emergency Medicine, Lancashire Teaching Hospitals NHS Foundation Trust, UK, kirsty.challen@lthtr.nhs.uk

Chris Fitzimmons, Consultant in Paediatric Emergency Medicine, Sheffield Children's NHS Foundation Trust, UK, c.fitzsimmons@nhs.net

Tim Harris, Professor of Emergency Medicine, Barts Health NHS Trust, UK, Tim.Harris@bartshealth.nhs.uk

Fiona Lecky, Clinical Professor in Emergency Medicine, University of Sheffield, UK, f.e.lecky@sheffield.ac.uk

Andrew Lee, Reader of Global Public Health, University of Sheffield, UK, andrew.lee@sheffield.ac.uk

Ian Maconochie, Consultant in Paediatric Emergency Medicine, Imperial College Healthcare NHS

Trust, UK, I.maconochie@ic.ac.uk

Darren Walter, Consultant in Emergency Medicine, Manchester University NHS Foundation Trust, UK,

darren.walter@manchester.ac.uk

Corresponding author:

Steve Goodacre, School of Health and Related Research (ScHARR), University of Sheffield, Regent

Court, Regent Street, Sheffield, S10 1UL

s.goodacre@sheffield.ac.uk

ORCID ID: 0000-0003-0803-8444

Word count: 3161

Abstract

Objectives

We aimed to derive and validate a triage tool, based on clinical assessment alone, for predicting adverse outcome in acutely ill adults with suspected COVID-19 infection.

Methods

We undertook a mixed prospective and retrospective observational cohort study in 70 emergency departments across the United Kingdom (UK). We collected presenting data from 22445 people attending with suspected COVID-19 between 26 March 2020 and 28 May 2020. The primary outcome was death or organ support (respiratory, cardiovascular, or renal) by record review at 30 days. We split the cohort into derivation and validation sets, developed a clinical score based on the coefficients from multivariable analysis using the derivation set, and the estimated discriminant performance using the validation set.

Results

We analysed 11773 derivation and 9118 validation cases. Multivariable analysis identified that age, sex, respiratory rate, systolic blood pressure, oxygen saturation/inspired oxygen ratio, performance status, history of renal impairment, and respiratory distress were retained in analyses restricted to the ten or fewer predictors. We used findings from multivariable analysis and clinical judgement to develop a score based on the NEWS2 score, age, sex, and performance status. This had a c-statistic of 0.80 (95% confidence interval 0.79-0.81) in the validation cohort and predicted adverse outcome with sensitivity 0.98 (0.97-0.98) and specificity 0.38 (0.38-0.39) for scores above four points.

Conclusion

A clinical score based on NEWS2, age, sex, and performance status predicts adverse outcome with good discrimination in adults with suspected COVID-19 and can be used to support decision-making in emergency care.

Registration

ISRCTN registry, ISRCTN28342533, http://www.isrctn.com/ISRCTN28342533

Introduction

The initial management of acutely ill people with suspected COVID-19 involves assessing the risk of
adverse outcome and the need for life-saving intervention, to then determine decisions around
hospital admission and inpatient referral.[1-5] Triage tools can assist decision-making by combining
information from clinical assessment in a structured manner to predict the risk of adverse outcome.
They can take the form of a score that increases with the predicted risk of adverse outcome or a rule
that categorises patients into groups according to their risk or their intended management. Inclusion
of laboratory and radiological information can improve prediction but requires hospital attendance,
increases emergency department (ED) length of stay, and increases the infection risk related to
repeated patient contacts. An appropriate triage tool for suspected COVID-19 needs to be based on
clinical assessment alone and applicable to people with suspected COVID-19.

We designed the Pandemic Influenza Triage in the Emergency Department (PAINTED) study
following the 2009 H1N1 influenza pandemic to develop and evaluate triage tools in any future
influenza pandemic.[6] We changed PAINTED to the Pandemic Respiratory Infection Emergency
System Triage (PRIEST) study in January 2020 to address any pandemic respiratory infection,
including COVID-19. The United Kingdom (UK) Department of Health and Social Care activated
PRIEST on 20 March 2020 to develop and evaluate triage tools in the COVID-19 pandemic. Initial
descriptive analysis of the PRIEST data showed that adults presenting to the ED with suspected
COVID-19 have much higher rates of COVID-19 positivity, hospital admission and adverse outcome
than children.[7] We therefore decided to undertake separate studies in adults and children, and
only develop a new triage tools in adults, which we present here.

Evaluation of existing triage tools using the PRIEST study data suggested that CURB-65, [8] the
National Early Warning Score version 2 (NEWS2) [9] and the Pandemic Modified Early Warning Score
(PMEWS) [10] provide reasonable prediction for adverse outcome in suspected COVID-19 (c-
statistics 0.75 to 0.77). [11] Scope therefore existed to develop a specific triage tool for COVID-19
with better prediction for adverse outcome.

Aims and objectives

We aimed to derive and validate a triage tool in the form of an illness severity score, based on
clinical assessment alone, for predicting adverse outcome in acutely ill adults with suspected COVID-
19 infection.

Methods

Design

We designed PRIEST as an observational study to collect standardised predictor variables recorded in
the ED, which we would then use to derive and validate new tools for predicting adverse outcome
up to 30 days after initial hospital presentation. The study did not involve any change to patient care.
Hospital admission and discharge decisions were made according to usual practice, informed by local
and national guidance.

Setting and population

We identified consecutive patients presenting to the ED of participating hospitals with suspected
COVID-19 infection. Patients were eligible if they met the clinical diagnostic criteria [12] of fever ($\geq$
37.8° C) and acute onset of persistent cough (with or without sputum), hoarseness, nasal discharge
or congestion, shortness of breath, sore throat, wheezing, or sneezing. This was determined on the
basis of the assessing clinician recording that the patient had suspected COVID-19 or completing a
standardised assessment form designed for suspected pandemic respiratory infection [6].

Interventions

For this study we planned to develop a triage tool in the form of an illness severity score based on
clinical assessment and routine observations that any health care professional could use to rapidly
estimate the risk of adverse outcome. The score would be based on a number of categorised
variables, with points allocated to each category of each variable, which would then be summed to
give a total score reflecting the predicted risk of adverse outcome. To enhance usability, we planned
to (a) use a restricted number of variables, rather than all potentially predictive variables, and (b)
categorise variables in accordance with currently used scores, unless there was clear evidence that
these categories provided suboptimal prediction.

Data collection

Data collection was both prospective and retrospective. Participating EDs were provided with a
standardised data collection form (Appendix 1) that included variables used in existing triage tools or
considered to be potentially useful predictors of adverse outcome. Participating sites could adapt
the form to their local circumstances, including integrating it into electronic or paper clinical records
to facilitate prospective data collection, or using it as a template for research staff to retrospectively
extract data from clinical records. We did not seek consent to collect data but information about the

study was provided in the ED and patients could withdraw their data at their request. Patients with
multiple presentations to hospital were only included once, using data from the first presentation
identified by research staff.

Outcome measurement

Research staff at participating hospitals reviewed patient records at 30 days after initial attendance
and recorded outcomes using the follow-up form in Appendix 2. The primary outcome was death or
major organ support (respiratory, cardiovascular, or renal) up to 30 days after initial attendance.
Death and major organ support were also analysed separately as secondary outcomes. Our primary
outcome definition reflected the need for triage tools to identify patients at risk of adverse outcome
or requiring life-saving intervention to prevent adverse outcome. Respiratory support was defined as
any intervention to protect the patient's airway or assist their ventilation, including non-invasive
ventilation or acute administration of continuous positive airway pressure. It did not include
supplemental oxygen alone or nebulised bronchodilators. Cardiovascular support was defined as any
intervention to maintain organ perfusion, such as inotropic drugs, or invasively monitor
cardiovascular status, such as central venous pressure or pulmonary artery pressure monitoring, or
arterial blood pressure monitoring. It did not include peripheral intravenous cannulation or fluid
administration. Renal support was defined as any intervention to assist renal function, such as
haemofiltration, haemodialysis, or peritoneal dialysis. It did not include intravenous fluid
administration.

Analysis

We randomly split the study population into derivation and validation cohorts by randomly
allocating the participating sites to one or other cohort. We developed a score based on the
prognostic value of predictor variables in multivariable analysis of the derivation cohort and expert
judgements regarding clinical usability. Candidate predictors were combined in a multivariable
regression with Least Absolute Shrinkage and Selection Operator (LASSO) using ten sample cross
validation to select the model. The LASSO begins with a full model of candidate predictors and
simultaneously performs predictor selection and penalisation during model development to avoid
overfitting. The LASSO was performed twice: once where the number of predictors were
unrestricted, and a second time when the LASSO was restricted to pick ten predictors. Fractional
polynomials were used to model non-linear relationships for continuous variables.

We excluded cases from all analyses if age or outcome data were missing. We undertook three multivariable analyses, using different approaches to missing predictor variable data in the derivation cohort: (1) Complete case; (2) Multiple imputation using chained equations; (3) Deterministic imputation with missing predictor data assumed to be normal, where applicable. We did not consider any predictor with more than 50% missing data across the cohort for inclusion in the predictive model.

Clinical members of the research team reviewed the models and selected variables for inclusion in the triage tool, based on their prognostic value in the model, the clinical credibility of their association with adverse outcome, and their availability in routine clinical care. We categorised continuous variables, using recognised categories from existing scores where appropriate, while checking that categorisation reflected the relationship between the variable and adverse outcome in the derivation data. We then assigned integer values to each category of predictor variable, taking into account the points allocated to the category in existing scores, and the coefficient derived from a multivariable logistic regression model using categorised continuous predictors. This generated a composite clinical score in which risk of adverse outcome increased with the total score.

We applied the clinical score to the validation cohort, calculating diagnostic parameters at each threshold of the score, constructing a receiver-operating characteristic (ROC) curve, calculating the area under the ROC curve (c-statistic) and calculating the proportion with an adverse outcome at each level of the score. We used deterministic imputation to handle missing data in the validation cohort, assuming missing predictor variable data were normal, but excluding cases with more than a pre-specified number of predictor variables missing. We also undertook a complete case sensitivity analysis.

The sample size was dependent on the size and severity of the pandemic, but based on a previous study in the 2009 H1N1 influenza pandemic we estimated we would need to collect data from 20,000 patients across 40-50 hospitals to identify 200 with an adverse outcome, giving sufficient power for model derivation. In the event, the adverse outcome rate in adults was much higher in the COVID-19 pandemic, giving us adequate power to undertake derivation and validation of triage tools to predict all three outcomes.

Patient and public involvement

The Sheffield Emergency Care Forum (SECF) is a public representative group interested in emergency care research. [13] Members of SECF advised on the development of the PRIEST study and two members joined the Study Steering Committee. Patients were not involved in the recruitment to and conduct of the study. We are unable to disseminate the findings to study participants directly.

Ethical approval

The North West - Haydock Research Ethics Committee gave a favourable opinion on the PAINTED study on 25 June 2012 (reference 12/NW/0303) and on the updated PRIEST study on 23rd March 2020. The Confidentiality Advisory Group of the Health Research Authority granted approval to collect data without patient consent in line with Section 251 of the National Health Service Act 2006.

Results

The PRIEST study recruited 22485 patients from 70 EDs across 53 sites between 26 March 2020 and 28 May 2020. We included 20889 in the analysis after excluding 39 who requested withdrawal of their data, 1530 children, 20 with missing outcome data, and seven with missing age. The derivation cohort included 11773 patients and the validation cohort 9118. Table 1 shows the characteristics of the derivation and validation cohorts.

Table 1: Characteristics of the study population (derivation and validation cohorts)

Characteristic	Statistic/level	Derivation	Validation
Age (years)	N	11773	9118
	Mean (SD)	62.4 (19.9)	62.4 (19.5)
	Median (IQR)	64 (48,79)	64 (48,79)
Sex	Missing	137	56
	Male	5746 (49.4%)	4455 (49.2%)
	Female	5890 (50.6%)	4607 (50.8%)
Ethnicity	Missing/prefer not to say	1819	2379
	UK/Irish/other white	8376 (84.1%)	5867 (87.1%)
	Asian	699 (7%)	345 (5.1%)
	Black/African/Caribbean	368 (3.7%)	272 (4%)
	Mixed/multiple ethnic groups	178 (1.8%)	69 (1%)
	Other	333 (3.3%)	186 (2.8%)
Presenting features	Cough	7248 (61.6%)	5737 (62.9%)
	Shortness of breath	8570 (72.8%)	7000 (76.8%)
	Fever	5714 (48.5%)	4562 (50%)
Comorbidities	Hypertension	3627 (30.8%)	2807 (30.8%)
	Heart Disease	2512 (21.3%)	2188 (24%)
	Diabetes	2394 (20.3%)	1735 (19%)

Characteristic	Statistic/level	Derivation	Validation
	Asthma	1867 (15.9%)	1541 (16.9%)
	Other chronic lung disease	2047 (17.4%)	1717 (18.8%)
	Renal impairment	1074 (9.1%)	856 (9.4%)
	Active malignancy	577 (4.9%)	543 (6%)
	Immunosuppression	312 (2.7%)	319 (3.5%)
	Steroid therapy	303 (2.6%)	254 (2.8%)
	No chronic disease	3385 (28.8%)	2406 (26.4%)
Symptom duration (days)	N	10790	8087
	Mean (SD)	8.1 (9.1)	7.6 (8.6)
	Median (IQR)	5 (2,10)	5 (2,10)
Heart rate (beats/min)	N	11506	8954
	Mean (SD)	94.7 (21.5)	95.2 (21.7)
	Median (IQR)	93 (80,108)	94 (80,109)
Respiratory rate (breaths/min)	N	11438	8908
	Mean (SD)	23.1 (6.9)	23.4 (7.1)
	Median (IQR)	22 (18,26)	22 (18,26)
Systolic BP (mmHg)	N	11423	8875
	Mean (SD)	134.5 (24.9)	134.8 (25)
	Median (IQR)	133 (118,149)	133 (118,150)
Diastolic BP (mmHg)	N	11373	8839
	Mean (SD)	78.3 (15.8)	78.2 (16.5)
	Median (IQR)	78 (68,88)	78 (68,88)
Temperature (°C)	N	11307	8924
	Mean (SD)	37.1 (1.1)	37.2 (1.1)
	Median (IQR)	37 (36.4,37.8)	37 (36.5,37.9)
Oxygen saturation (%)	N	11658	8974
	Mean (SD)	94.9 (6.2)	94.4 (7.5)
	Median (IQR)	96 (94,98)	96 (94,98)
Air or supplementary oxygen	Missing	4113	4735
	On air	5243 (68.4%)	2544 (58%)
	On supplementary oxygen	2417 (31.6%)	1839 (42%)
Supplementary inspired oxygen (%)	N	2417	1839
	Mean (SD)	45.9 (21.9)	48.6 (22.5)
	Median (IQR)	36 (28,60)	36 (28,80)
Glasgow Coma Scale	N	8627	6801
	Mean (SD)	14.6 (1.4)	14.6 (1.4)
	Median (IQR)	15 (15,15)	15 (15,15)
Consciousness	Missing	1515	872
	Alert	9774 (95.3%)	7794 (94.5%)
	Verbal	333 (3.2%)	307 (3.7%)
	Pain	101 (1%)	82 (1%)
	Unresponsive	50 (0.5%)	63 (0.8%)
Performance status	Missing	620	458

Characteristic	Statistic/level	Derivation	Validation
	1. Unrestricted normal activity	5989 (53.7%)	4547 (52.5%)
	2. Limited strenuous activity, can do light activity	1315 (11.8%)	1056 (12.2%)
	3. Limited activity, can self-care	1565 (14%)	1211 (14%)
	4. Limited self-care	1494 (13.4%)	1155 (13.3%)
	5. Bed/chair bound, no self-care	790 (7.1%)	691 (8%)
Admitted at initial assessment	Missing	7	21
	No	3744 (31.8%)	3122 (34.3%)
	Yes	8022 (68.2%)	5975 (65.7%)
Respiratory pathogen	COVID-19	3660 (31.1%)	2861 (31.4%)
	Influenza	2 (0%)	25 (0.3%)
	Other	912 (7.7%)	809 (8.9%)
	None identified	7199 (61.1%)	5423 (59.5%)
Mortality status	Missing	0	3
	Alive	9949 (84.5%)	7619 (83.6%)
	Dead	1824 (15.5%)	1496 (16.4%)
	Death with organ support*	335 (18.4%)	371 (24.8%)
	Death with no organ support*	1489 (81.6%)	1125 (75.2%)
Organ support	Respiratory	939 (8%)	1005 (11%)
	Cardiovascular	316 (2.7%)	201 (2.2%)
	Renal	104 (0.9%)	114 (1.3%)
	Any	999 (8.5%)	1059 (11.6%)

Table 2 shows summary statistics for each predictor variable in those with and without adverse
outcome in the derivation sample, and univariate odds ratios for prediction of adverse outcome.
Physiological variables were categorised to reflect their expected relationships with adverse
outcome.

Table 2: Univariate analysis of predictor variables for each adverse outcome
definition (derivation cohort)

Predictor Variable	Category (categorical variables)	n (outcome)		Odds Ratio	P value	95% CI
		Adverse	Non adverse			
Age (n=11773)				1.04	0.00	(1.04, 1.04)
Ethnicity Category (n=9954)	Ref=UK/Irish/other white	1800	6576			
	Asian	138	561	0.90	0.28	(0.74, 1.09)
	Black/African/Caribbean	72	296	0.89	0.38	(0.68, 1.16)
	Mixed/multiple ethnic groups	29	149	0.71	0.10	(0.48, 1.06)

	Other	39	294	0.48	0.00	(0.35, 0.68)
Shortness of breath (n=11746)	Ref=No	552	2624			
	Yes	1924	6646	1.38	0.00	(1.24, 1.53)
Cough (n=11746)	Ref=No	1085	3413			
	Yes	1391	5857	0.75	0.00	(0.68, 0.82)
Fever (n=11746)	Ref=No	1298	4734			
	Yes	1178	4536	0.95	0.23	(0.87, 1.04)
Hypertension (n=11732)	Ref=No	1470	6635			
	Yes	1014	2613	1.75	0.00	(1.60, 1.92)
Heart Disease (n=11732)	Ref=No	1705	7515			
	Yes	779	1733	1.98	0.00	(1.79, 2.19)
Diabetes (n=11732)	Ref=No	1764	7574			
	Yes	720	1674	1.85	0.00	(1.67, 2.04)
Asthma (n=11732)	Ref=No	2182	7683			
	Yes	302	1565	0.68	0.00	(0.60, 0.78)
Other chronic lung disease (n=11732)	Ref=No	1959	7726			
	Yes	525	1522	1.36	0.00	(1.22, 1.52)
Renal impairment (n=11732)	Ref=No	2085	8573			
	Yes	399	675	2.43	0.00	(2.13, 2.78)
Active malignancy (n=11732)	Ref=No	2288	8867			
	Yes	196	381	1.99	0.00	(1.67, 2.38)
Immunosuppression (n=11732)	Ref=No	2403	9017			
	Yes	81	231	1.32	0.04	(1.02, 1.70)
Steroid therapy (n=11732)	Ref=No	2406	9023			
	Yes	78	225	1.30	0.05	(1.00, 1.69)
Symptom duration (n=10790)				0.97	0.00	(0.96, 0.97)
Number current medications (n=11183)				1.09	0.00	(1.08, 1.10)

Respiratory rate (n=11773)	Ref=12-20 or missing	666	5039			
	<9	3	5	4.54	0.04	(1.08, 19.04)
	9-11	3	8	2.84	0.12	(0.75, 10.72)
	21-24	592	2180	2.05	0.00	(1.82, 2.32)
	>24	1224	2053	4.51	0.00	(4.05, 5.02)
Systolic Blood Pressure (n=11773)	Ref=111-219 or missing	1894	8059			
	101-110	273	741	1.57	0.00	(1.35, 1.82)
	91-100	175	315	2.36	0.00	(1.95, 2.86)
	<91	138	142	4.14	0.00	(3.25, 5.25)
	>219	8	28	1.22	0.63	(0.55, 2.67)
Heart rate (n=11773)	Ref=51-90 or missing	1025	4335			
	<41	15	42	1.51	0.17	(0.83, 2.73)
	41-50	12	42	1.21	0.57	(0.63, 2.30)
	91-110	793	3095	1.08	0.13	(0.98, 1.20)
	111-130	455	1362	1.41	0.00	(1.25, 1.60)
	>130	188	409	1.94	0.00	(1.61, 2.34)
Temperature (n=11773)	Ref=36.1-38.0 or missing	1522	6723			
	35.1-36	255	948	1.19	0.02	(1.02, 1.38)
	38.1-39	452	1131	1.77	0.00	(1.56, 2.00)
	>39.0	168	384	1.93	0.00	(1.60, 2.34)
	<35.1	91	99	4.06	0.00	(3.04, 5.43)
GCS Total (n=8627)	Ref=13-15	1582	6587			
	9-12	189	148	5.32	0.00	(4.26, 6.64)
	<=8	74	47	6.56	0.00	(4.53, 9.49)
AVPU (n=10258)	Ref=Alert	1793	7981			
	Verbal	177	156	5.05	0.00	(4.05, 6.30)
	Pain	63	38	7.38	0.00	(4.92, 11.07)
	Unresponsive	32	18	7.91	0.00	(4.43, 14.13)
Performance status (n=11153)	Ref=Unrestricted normal activity	714	5275			
	Limited strenuous activity, can do light activity	272	1043	1.93	0.00	(1.65, 2.25)
	Limited activity, can self-care	443	1122	2.92	0.00	(2.55, 3.34)
	Limited self-care	574	920	4.61	0.00	(4.05, 5.25)
	Bed/chair bound, no self-care	340	450	5.58	0.00	(4.75, 6.56)
Severe respiratory distress (n=11773)	Ref=No	2292	9142			

	Yes	196	143	5.47	0.00	(4.39, 6.81)
Respiratory exhaustion (n=11773)	Ref=No	2404	9183			
	Yes	84	102	3.15	0.00	(2.35, 4.21)
Severe dehydration (n=11773)	Ref=No	2416	9197			
	Yes	72	88	3.11	0.00	(2.27, 4.27)
Previous attendance (n=11773)	Ref=No	2200	8389			
	Yes	288	896	1.23	0.01	(1.06, 1.41)
Known contact with Covid-19 case (n=1177)	Ref=No	2215	8434			
	Yes	273	851	1.22	0.01	(1.06, 1.41)
Central capillary refill (n=2935)	Ref=Normal	494	2171			
	Abnormal	102	168	2.67	0.00	(2.05, 3.48)

Supplementary Tables 1 to 3 show the results of multivariable analysis using complete case analysis, multiple imputation and deterministic imputation. Unrestricted LASSO on multiply imputed data included more predictors, with a higher c-statistic for the model (0.85), than the LASSO on deterministically imputed data or complete cases (c-statistics both 0.83). When restricted, there were eight predictors that were retained by LASSO in all three analyses (age, sex, respiratory rate, systolic BP, oxygen saturation/inspired oxygen ratio, history of renal impairment, performance status, and respiratory distress). C-statistics for the restricted models using deterministic imputation and complete case analysis (0.82) were slightly lower than c-statistics for the respective unrestricted models (0.83).

We developed a score through the following steps:
1. Clinical review judged that the eight predictors are clinically credible; that age, sex, respiratory rate, systolic BP, oxygen saturation and inspired oxygen are routinely recorded in administrative systems and early warning scores (although the ratio of oxygen saturation to inspired oxygen is not routinely recorded); and that many EDs routinely record a measure of performance status for suspected COVID-19 cases that could be mapped onto our scale.
2. We decided to include consciousness, temperature and heart rate, as these are routinely recorded alongside other physiological variables in early warning scores, and added prognostic value in the full models.

3. We created categories for age based on the observed multivariate association between age
    and outcome in our data, and categories for respiratory rate, heart rate, oxygen saturation,
    inspired oxygen, systolic BP, consciousness and temperature based on those used in the
    NEWS2 early warning score.

4. We created a multivariable logistic regression model using categorised predictor variables
    (Supplementary Table 4) and compared the coefficients for each category of predictor
    variable in the NEWS2 score to the points allocated in the NEWS2 score. We judged that the
    inconsistencies between the coefficients and the points used in NEWS2 were insufficient to
    justify allocating alternative points in our score. We allocated points to categories of age, sex,
    performance status, renal history, and respiratory distress, based on the coefficients in the
    model.

5. We removed renal history and respiratory distress from the multivariable model
    (Supplementary Table 5), noted that this made no meaningful difference to the c-statistic
    (0.82 in both models) and, given concerns about subjectivity and lack of routine recording,
    decided not to include them in the score.

The developed score is shown in Figure 1. We applied the score to the validation cohort. Figure 2
shows the ROC curve, with a c-statistic of 0.80 (95% CI 0.79 to 0.81) for the score. Sensitivity analysis
using complete cases only gave a c-statistic of 0.79 (95% CI 0.77 to 0.80). Supplementary Figures 1
and 2 show the calibration plots for the unrestricted and restricted LASSO models applied to the
validation cohort. The c-statistics (0.82 and 0.81 respectively, compared with 0.80 for the score)
indicate the effect of restricting the number of variables and then developing a score had upon
discrimination. Figure 3 shows the probability of adverse outcome for each value of the score. Table
3 shows the sensitivity and specificity for predicting outcome at each threshold of the triage tool.

Table 3: Sensitivity, specificity, PPV, NPV and proportion with a positive score at each score threshold for predicting the primary outcome of death or organ support, validation cohort

Score threshold	Proportion with positive score	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
>0	0.97	1.00 (1.00, 1.00)	0.04 (0.03, 0.04)	0.25 (0.24, 0.25)	0.99 (0.98, 1.00)
>1	0.92	1.00 (1.00, 1.00)	0.10 (0.10, 0.10)	0.26 (0.26, 0.26)	0.99 (0.99, 0.99)
>2	0.87	0.99 (0.99, 0.99)	0.17 (0.17, 0.18)	0.28 (0.27, 0.28)	0.99 (0.98, 0.99)
>3	0.80	0.99 (0.98, 0.99)	0.26 (0.26, 0.27)	0.30 (0.29, 0.30)	0.98 (0.98, 0.99)
>4	0.73	0.98 (0.97, 0.98)	0.34 (0.34, 0.35)	0.32 (0.32, 0.33)	0.98 (0.98, 0.98)
>5	0.66	0.95 (0.95, 0.95)	0.43 (0.43, 0.44)	0.35 (0.34, 0.35)	0.96 (0.96, 0.97)
>6	0.59	0.91 (0.91, 0.91)	0.51 (0.50, 0.51)	0.37 (0.36, 0.37)	0.95 (0.94, 0.95)

>7	0.53	0.86 (0.85, 0.86)	0.58 (0.58, 0.58)	0.39 (0.39, 0.40)	0.93 (0.93, 0.93)
>8	0.46	0.80 (0.79, 0.80)	0.65 (0.64, 0.65)	0.42 (0.41, 0.42)	0.91 (0.91, 0.91)
>9	0.40	0.73 (0.72, 0.74)	0.71 (0.71, 0.71)	0.44 (0.44, 0.45)	0.89 (0.89, 0.89)
>10	0.33	0.65 (0.64, 0.66)	0.77 (0.77, 0.78)	0.48 (0.47, 0.48)	0.87 (0.87, 0.88)
>11	0.27	0.57 (0.56, 0.58)	0.82 (0.82, 0.82)	0.50 (0.50, 0.51)	0.86 (0.85, 0.86)
>12	0.21	0.47 (0.46, 0.48)	0.87 (0.87, 0.88)	0.54 (0.53, 0.55)	0.84 (0.84, 0.84)
>13	0.16	0.37 (0.37, 0.38)	0.91 (0.91, 0.91)	0.56 (0.56, 0.57)	0.82 (0.82, 0.82)
>14	0.12	0.29 (0.28, 0.30)	0.94 (0.94, 0.94)	0.60 (0.58, 0.61)	0.81 (0.80, 0.81)
>15	0.09	0.23 (0.22, 0.23)	0.96 (0.96, 0.96)	0.63 (0.62, 0.64)	0.80 (0.79, 0.80)
>16	0.06	0.17 (0.17, 0.18)	0.97 (0.97, 0.97)	0.66 (0.64, 0.67)	0.79 (0.78, 0.79)
>17	0.04	0.12 (0.11, 0.12)	0.98 (0.98, 0.98)	0.69 (0.67, 0.70)	0.78 (0.77, 0.78)
>18	0.03	0.08 (0.07, 0.08)	0.99 (0.99, 0.99)	0.67 (0.65, 0.70)	0.77 (0.77, 0.77)
>19	0.02	0.05 (0.04, 0.05)	0.99 (0.99, 1.00)	0.74 (0.71, 0.76)	0.77 (0.76, 0.77)
>20	0.01	0.03 (0.03, 0.03)	1.00 (1.00, 1.00)	0.77 (0.73, 0.80)	0.76 (0.76, 0.77)
>21	0.01	0.02 (0.02, 0.02)	1.00 (1.00, 1.00)	0.81 (0.76, 0.85)	0.76 (0.76, 0.77)
>22	0.00	0.01 (0.01, 0.01)	1.00 (1.00, 1.00)	0.84 (0.76, 0.90)	0.76 (0.76, 0.76)
>23	0.00	0.01 (0.00, 0.01)	1.00 (1.00, 1.00)	0.87 (0.76, 0.94)	0.76 (0.76, 0.76)
>24	0.00	0.00 (0.00, 0.00)	1.00 (1.00, 1.00)	0.86 (0.66, 0.96)	0.76 (0.76, 0.76)

Supplementary Figures 3 and 4 show the ROC curves, and Supplementary Tables 6 and 7 show the predictive performance of the score when applied to the secondary outcomes of organ support and death without organ support in the validation cohort. The score provided better prognostic discrimination for death without organ support (c-statistic 0.83, 95% CI 0.82 to 0.84) than for organ support (0.68, 95% CI 0.67 to 0.69).

Discussion

We have developed a clinical illness severity score for acutely ill patients with suspected COVID-19 that combines the NEWS2 score, age, sex, and performance status to predict the risk of death or receipt of organ support. The score ranges from zero to 29 points, with a score greater than four predicting adverse outcome with high sensitivity. In developing the score, we tried to optimise usability without compromising performance. Usability was optimised by basing the score on the existing NEWS2 score and only adding easily available information. The c-statistic of the score on the validation cohort was 0.80, compared with 0.82 and 0.81 when the unrestricted and restricted models were applied to the validation cohort, suggesting that simplifying the tool did not excessively compromise prediction.

We previously analysed the performance of triage tools that have been recommended in guidelines for the initial assessment of acutely ill people with sepsis, and showed that CURB-65 (c-statistic 0.75),

PMEWS (0.77) and NEWS2 (0.77) offer good prediction for adverse outcome. Our new triage tool offers improved prediction with the addition of three variables.

Previous research

A living systematic review [14] has identified 50 prognostic models for adverse outcome in people with diagnosed COVID-19. C-statistics ranged from 0.68 to 0.99, and the most frequently used predictor variables were age, sex, comorbidities, temperature, lymphocyte count, C reactive protein, creatinine, and imaging features. Recently the ISARIC WHO Clinical Characterisation Protocol developed and validated the 4C Mortality Score [15] that predicts the mortality risk for people admitted with COVID-19 with better discriminant performance than 15 pre-existing risk stratification scores (c-statistic 0.77 versus 0.61-0.76).

Mortality prediction scores have an important role predicting mortality in hospital admissions but have limitations as triage tools. The inclusion of laboratory data as predictor variables usually requires hospital attendance, prolongs ED stay and prevents the rapid assessment required in ED or prehospital settings. Furthermore triage tools need to predict need for life-saving intervention rather than just mortality, and ideally need to be developed and evaluated in a relevant cohort, i.e. those with suspected COVID-19, including those not admitted to hospital after assessment.

Rapid clinical scores have been proposed or evaluated in several studies. Liao *et al* [16] proposed adding age>65 years to the NEWS2 score to aid decision-making, based on early experience of the pandemic in China. Myrstad *et al* [17] reported a c-statistic of 0.822 (95% CI 0.690 to 0.953) for NEWS2 predicting death or severe disease in a small study (N=66) of people hospitalised with confirmed COVID-19. Hu *et al* [18] reported c-statistics of 0.833 (0.737 to 0.928) for the Rapid Emergency Medicine Score (REMS) and 0.677 (0.541 to 0.813) for the Modified Emergency Medicine Score (MEWS) for predicting mortality in critically ill patients with COVID-19. Haimovich *et al* [19] developed the quick COVID-19 severity index, consisting of respiratory rate, oxygen saturation, and oxygen flow rate, which predicted respiratory failure within 24 hours in adults admitted with COVID-19 requiring supplemental oxygen with a c-statistic of 0.81 (0.73 to 0.89).

Strengths and limitations

We collected data from a clinically relevant population of patients presenting with suspected COVID-19 across a varied range of EDs. The large sample size and high rate of adverse outcome provided good statistical power to support analysis of many predictor variables in multivariable analysis and

allowed us to estimate parameters with a high degree of precision. An important limitation is that
retrospective data collection resulted in some missing and may have resulted in some inaccuracy of
predictor variable recording. Recording of inspired oxygen concentration was subject to a
particularly high rate of missing data. We anticipated this problem and pre-specified analyses
involving multiple imputation, deterministic imputation, and complete case analysis to explore the
impact of missing data. There was reasonable concordance between the models. Another potential
limitation is that we may have missed adverse outcomes if patients attended a different hospital
after initial hospital discharge. This is arguably less likely in the context of a pandemic, in which
movements between regions were curtailed, but cannot be discounted. The 5-point scale we used
for determining performance status has not been widely used or evaluated, although the 9-point
clinical frailty index maps onto it reasonably well. Finally, although our triage tool can be used in the
prehospital or community setting, we recommend caution in extrapolating our findings to settings
where there is likely to be a lower prevalence of adverse outcome.

Implications for practice
Our clinical score could be used to support ED decision-making around hospital admission and
inpatient referral. A score greater than four could provide an appropriate balance of sensitivity and
specificity to support hospital admission decisions, while a higher threshold could be used to select
patients for critical care referral. However, triage tools should support and not replace clinical
decision-making, and patient preferences and values must be considered. Our triage tool could also
be used to support prehospital and community decision-making around decisions to refer for
hospital assessment. However, further validation is required to determine the performance of the
tool in these settings.

In summary, we have developed a clinical score that can provide a rapid and accurate assessment of
the risk of adverse outcome in adults who are acutely ill with suspected COVID-19.

Competing interests
All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf
and declare: grant funding to their employing institutions from the National Institute for Health
Research; no financial relationships with any organisations that might have an interest in the
submitted work in the previous three years; no other relationships or activities that could appear to
have influenced the submitted work.

References

1. World Health Organisation. Clinical care of severe acute respiratory infections – Tool kit.
   https://www.who.int/publications-detail/clinical-care-of-severe-acute-respiratory-
   infections-tool-kit (accessed 28/04/2020)
2. International Federation for Emergency Medicine. Global Response to COVID-19 for
   Emergency Healthcare Systems and Providers: From the IFEM Task Force on ED Crowding
   and Access Block. https://www.ifem.cc/coronavirus-2019-information/ (accessed
   15/06/2020)

3. NHS. Clinical guide for the management of emergency department patients during the coronavirus pandemic. 17 March 2020 Version 1 https://www.england.nhs.uk/coronavirus/secondary-care/other-resources/specialty-guides/#ae (accessed 15/06/2020)

4. National Institute for Health and Care Excellence. COVID-19 rapid guideline: managing suspected or confirmed pneumonia in adults in the community. Published: 3 April 2020. www.nice.org.uk/guidance/ng165 (accessed 28/04/2020)

5. American College of Emergency Physicians. Guide to Coronavirus Disease (COVID-19) https://www.acep.org/corona/covid-19-field-guide/cover-page/

6. Goodacre S, Irving A, Wilson R, Beever D, Challen K. The PAndemic INfluenza Triage in the Emergency Department (PAINTED) pilot cohort study. Health Technol Assess 2015;19(3):1-70.

7. Goodacre S, Thomas B, Lee E, et al. Characterisation of 22446 patients attending UK emergency departments with suspected COVID-19 infection: Observational cohort study [Preprint]. medRxiv 2020.08.10.20171496 https://www.medrxiv.org/content/10.1101/2020.08.10.20171496v1

8. Lim W, van der Eerden MM, Laing R, Boersma W, Karalus N, Town G, Lewis S, Macfarlane J. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. Thorax 2003; 58(5): 377–382.

9. Royal College of Physicians. (2017). National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. Updated report of a working party. London: RCP.

10. Challen K, Bright J, Bentley A, Walter D. Physiological-social score (PMEWS) vs. CURB-65 to triage pandemic influenza: a comparative validation study using community-acquired pneumonia as a proxy. BMC Health Serv Res. 2007; 7: 33.

11. Thomas B, Biggs K Goodacre S, et al. Prognostic accuracy of emergency department triage tools for adults with suspected COVID-19: The PRIEST observational cohort study. [Prerpint] medRxiv 2020.09.01.20185793; https://www.medrxiv.org/content/10.1101/2020.09.02.20185892v1

12. Public Health England. COVID-19: investigation and initial clinical management of possible cases. https://www.gov.uk/government/publications/wuhan-novel-coronavirus-initial-investigation-of-possible-cases/investigation-and-initial-clinical-management-of-possible-cases-of-wuhan-novel-coronavirus-wn-cov-infection#criteria (accessed 27/04/2020)

13. Hirst E, Irving A, Goodacre S. Patient and public involvement in emergency care research. Emerg Med J 2016;33:665–670.

14. Wynants Laure, Van Calster Ben, Collins Gary S, Riley Richard D, Heinze Georg, Schuit Ewoud et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal BMJ 2020; 369 :m1328

15. Knight Stephen R, Ho Antonia, Pius Riinu, Buchan Iain, Carson Gail, Drake Thomas M 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

16. Liao X, Wang B Kang Y. Novel coronavirus infection during the 2019–2020 epidemic: preparing intensive care units—the experience in Sichuan Province, China. Intensive Care Med 46, 357–360 (2020).

17. Myrstad M, Ihle-Hansen H, Tveita AA, Andersen EL, Nygård S, Tveit A, Berge T. National Early Warning Score 2 (NEWS2) on admission predicts severe disease and inhospital mortality from Covid-19 – a prospective cohort study. Scandinavian Journal of Trauma, Resuscitation and Emergency Medicine 2020; 28:66.

18. Hu H, Yao N, Qiu Y. Comparing Rapid Scoring Systems in Mortality Prediction of Critically Ill Patients With Novel Coronavirus Disease. Acad Emerg Med. 2020;27(6):461-468.

19. Haimovich A, Ravindra NG, Stoytchev S, Young HP, PerryWilson F, van Dijk D, Schulz WL, Taylor RA, Development and validation of the quick COVID-19 severity index (qCSI): a prognostic tool for early clinical decompensation Annals of Emergency Medicine (2020), doi: https://doi.org/10.1016/j.annemergmed.2020.07.022.

Figure 1: The PRIEST COVID-19 clinical severity score

Variable	Range	Score
Respiratory rate (per minute)	12-20	0
	9-11	1
	21-24	2
	<9 or >24	3
Oxygen saturation (%)	>95	0
	94-95	1
	92-93	2
	<92	3
Heart rate (per minute)	51-90	0
	41-50 or 91-110	1
	111-130	2
	<41 or >130	3
Systolic BP (mmHg)	111-219	0
	101-110	1
	91-100	2
	<91 or >219	3
Temperature (°C)	36.1-38.0	0
	35.1-36.0 or 38.1-39.0	1
	>39.0	2
	<35.1	3
Alertness	Alert	0
	Confused or not alert	3
Inspired oxygen	Air	0
	Supplemental oxygen	2
Sex	Female	0
	Male	1
Age (years)	16-49	0
	50-65	2
	66-80	3
	>80	4
Performance status	Unrestricted normal activity	0
	Limited strenuous activity, can do light activity	1
	Limited activity, can self-care	2
	Limited self-care	3
	Bed/chair bound, no self-care	4

Figure 2: ROC curve for the tool predicting the primary outcome of death or organ support, validation cohort
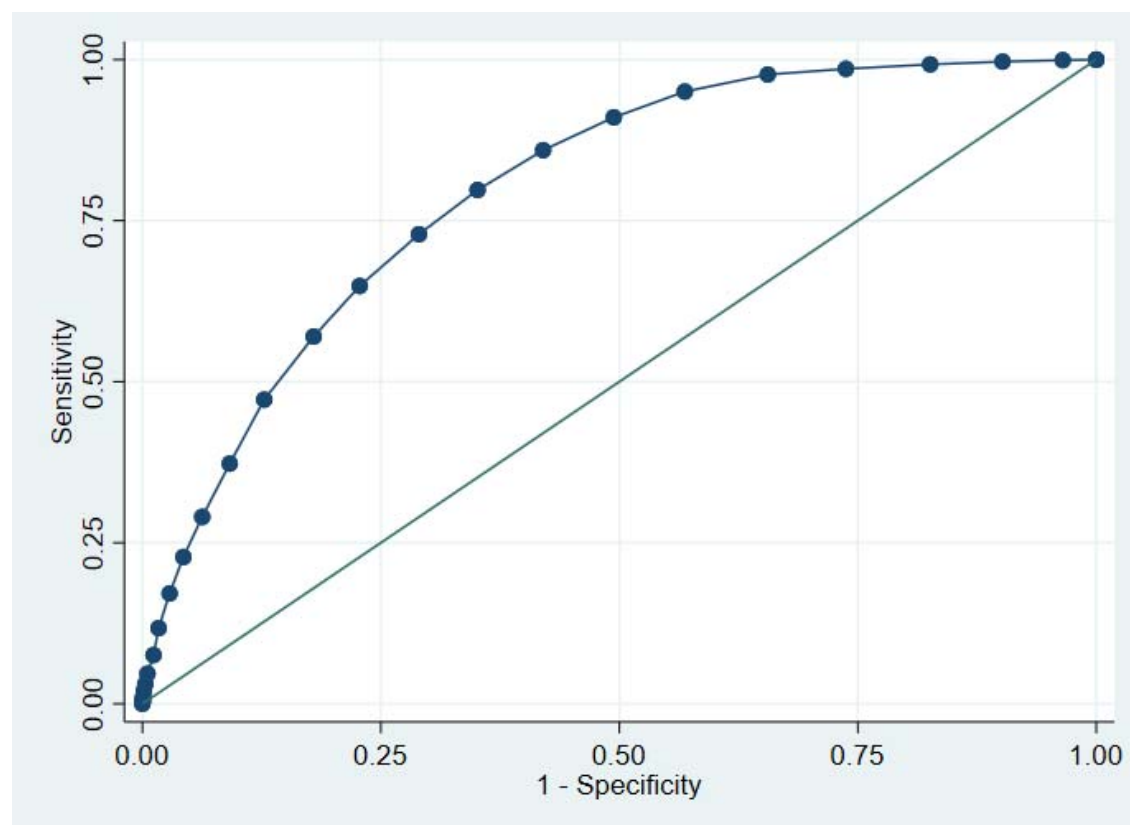
C-statistic 0.80 (95% CI 0.79-0.81)

Figure 3: Probability of adverse outcome for each value of the score, validation cohort
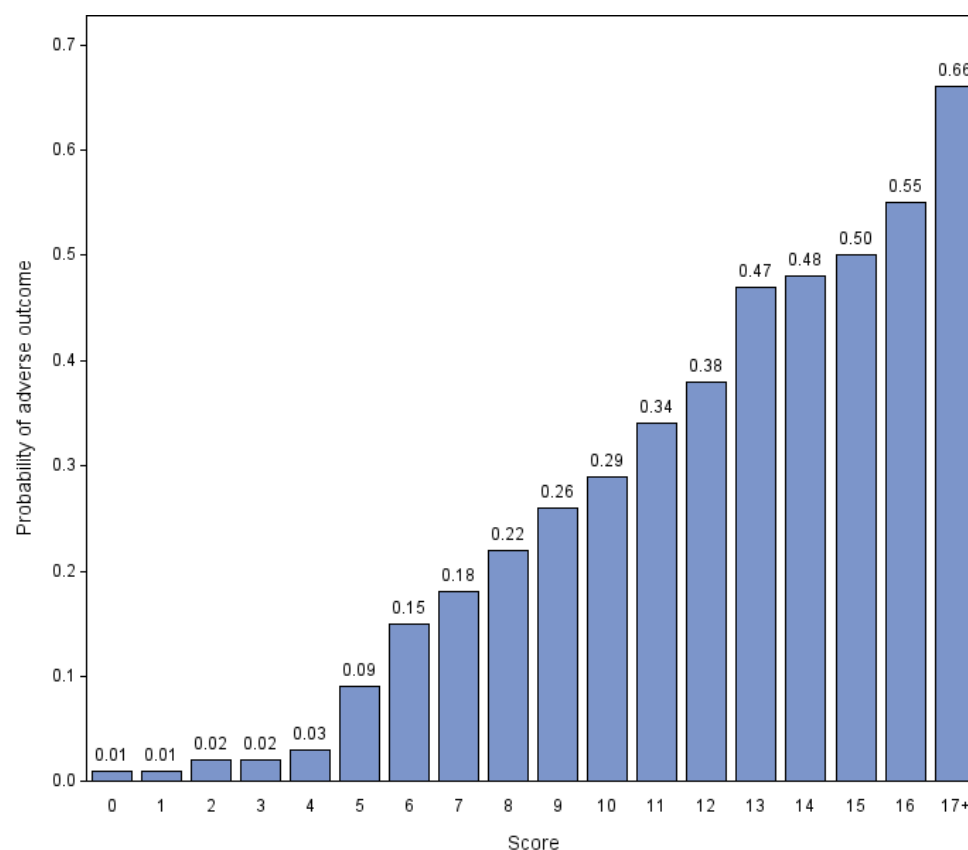
Supplementary Table 1: Multivariable analysis, complete case (N=5988)

Complete case

Lasso variable selection (unrestricted)

C-statistic: 0.83

Lasso variable selection (restricted to 10):

C-statistic: 0.82

Parameter		Coefficient		Parameter	Coefficient	
		Unstandardised	Standardised		Unstandardised	Standardised
Age		0.022	0.433	Age	0.015	0.303
Symptom duration		-0.000	-0.001	ln(respiratory rate)	0.840	0.225
ln(respiratory rate)		1.108	0.296	Systolic BP^-2	5228.37	0.131
Heart rate		0.002	0.039	Oxygen saturation/inspired ratio	-0.005	-0.580
ln(temperature)*temperature^3		0.000	0.076	Medication count	0.004	0.020
Systolic BP^-2		10578.93	0.266	Male sex	0.021	0.011
Oxygen saturation/inspired ratio		-0.006	-0.629	Renal impairment	0.065	0.019
Medication count		0.018	0.081	Performance status level 1	-0.084	-0.042
Male sex		0.342	0.171	Respiratory distress	0.090	0.017
Shortness of breath		0.132	0.057	Consciousness alert	-0.405	-0.104
Renal impairment		0.337	0.097	Constant	-2.725	
Diabetes		0.159	0.065			
Active malignancy		0.027	0.006			
Other chronic lung disease		-0.071	-0.027			
Clinically obese		0.240	0.063			
Tobacco or vape user		-0.203	-0.062			
Covid contact		0.227	0.069			
Performance status						
	1	-0.193	-0.096			
	2	-0.276	-0.090			
	4	0.042	0.015			
Respiratory distress		0.501	0.092			
Respiratory exhaustion		0.256	0.035			
Dehydration		0.684	0.086			
Consciousness						
	Alert	-0.551	-0.142			
	Pain	0.530	0.057			
Constant		-3.018				

Supplementary Table 2: Multivariable analysis, using multiple imputation (N=11636)

Multiple Imputation - 50 imputed data sets

Lasso variable selection (unrestricted)

C-statistic: 0.85

Lasso variable selection (restricted to 10):

C-statistic: 0.85

Parameter	Average coefficient		Number of times selected	Parameter	Average coefficient		Number of times selected
	Unstandardised	Standardised			Unstandardised	Standardised	
Age	0.022	0.446	50	Age	0.018	0.349	50
Symptom duration	-0.007	-0.060	50	ln(respiratory rate)	1.024	0.270	50
ln(respiratory rate)	1.276	0.337	50	systolic BP^-2	4484.729	0.113	50
Heart rate	0.002	0.042	50	Oxygen saturation/inspired ratio	-0.005	-0.602	50
ln(temp)*temp^3	1.17E-06	0.021	50	Male sex	0.035	0.017	50
Systolic BP^-2	8714.823	0.220	50	Renal impairment	0.081	0.023	50
Oxygen saturation/inspired ratio	-0.006	-0.643	50	Performance status level 1	-0.199	-0.099	50
Medication count	0.007	0.032	50	Performance status level 4	0.057	0.020	50
Male sex	0.312	0.156	50	Respiratory distress	0.166	0.028	50
Shortness of breath	0.187	0.083	50	Alert	-0.507	-0.133	50
Previous attendance	0.054	0.016	50				
Renal impairment	0.317	0.091	50	*Performance status level 5*	*0.019*	*0.005*	*48*
Asthma	-0.111	-0.040	50	*Diabetes*	*0.009*	*0.004*	*8*
Diabetes	0.175	0.070	50	*Medication count*	*0.000*	*0.000*	*2*
Active malignancy	0.302	0.065	50				
				Constant	-3.147		
Immunosuppression	0.285	0.046	50				
Other chronic lung disease	-0.116	-0.044	50				
Hypertension	0.063	0.029	50				
Clinically obese	0.278	0.067	50				
Tobacco or vape user	-0.248	-0.076	50				
Covid contact	0.184	0.054	50				
Performance status							
1	-0.237	-0.118	50				
2	-0.185	-0.060	50				
4	0.227	0.077	50				
5	0.234	0.060	50				
Respiratory distress	0.503	0.084	50				
Respiratory exhaustion	0.418	0.052	50				
Dehydration	0.480	0.056	50				
Consciousness							
Alert	-0.605	-0.159	50				

Pain	0.335	0.035	50
Unresponsive	0.408	0.031	50
Heart disease	*-0.033*	*-0.013*	*47*
Cough	*0.014*	*0.007*	*40*
Steroid use	*0.004*	*-0.001*	*5*
Pregnant	*0.032*	*-0.002*	*4*
Fever	*0.008*	*-0.004*	*1*
Constant	-2.405		

Supplementary Table 3: Multivariable analysis, using deterministic imputation (N=9891)

Deterministic imputation					
Lasso variable selection (unrestricted) C-statistic: 0.83			**Lasso variable selection (restricted to 10):** C-statistic: 0.82		

Parameter	Coefficient		Parameter	Coefficient	
	Unstandardised	Standardised		Unstandardised	Standardised
Age	0.024	0.487	Age	0.018	0.367
Symptom duration	-0.005	-0.049	ln(respiratory rate)	1.168	0.304
ln(respiratory rate)	1.448	0.377	Systolic BP^-2	5301.86	0.130
Heart rate	0.003	0.066	Oxygen saturation/inspired ratio	-0.005	-0.480
Systolic BP^-2	10326.8	0.254	Male sex	0.056	0.028
Oxygen saturation/inspired ratio	-0.005	-0.516	Renal impairment	0.118	-0.033
Male sex	0.355	0.178	Performance status level 1	-0.144	-0.071
Shortness of breath	0.218	0.096	Performance status level 4	0.042	0.014
Renal impairment	0.379	0.108	Respiratory distress	0.061	0.010
Asthma	-0.129	-0.048	Consciousness alert	-0.456	-0.111
Diabetes	0.212	0.085	Constant	-3.911	
Active malignancy	0.349	0.074			
Immunosuppression	0.295	0.048			
Other lung disease	-0.084	-0.032			
Hypertension	0.067	0.031			
Clinically obese	0.372	0.091			
Tobacco or vape user	-0.239	-0.074			
Covid contact	0.210	0.062			
Performance status					
1	-0.191	-0.095			
2	-0.187	-0.059			
4	0.211	0.069			
5	0.115	0.027			
Respiratory distress	0.408	0.069			
Respiratory exhaustion	0.464	0.059			
Dehydration	0.455	0.052			
Consciousness					
Alert	-0.646	-0.157			
Pain	0.446	0.044			
Unresponsive	0.115	0.007			
Constant	-2.882				

Supplementary Table 4: Logistic regression model based on selected categorised predictor variables

Predictor	Score allocated	Coefficient	SE	z	P-value	Upper 95% CI	Lower 95% CI
Respiratory rate							
12-20*	0						
9-11	1	0.534	1.394	0.380	0.702	-2.198	3.266
21-24	2	0.306	0.089	3.460	0.001	0.133	0.480
<9 or >24	3	0.657	0.082	8.020	0.000	0.496	0.817
Oxygen saturation							
>95%*	0						
94-95%	1	0.341	0.089	3.820	0.000	0.166	0.515
92-93%	2	0.659	0.111	5.910	0.000	0.441	0.878
<92%	3	1.095	0.085	12.900	0.000	0.929	1.261
Heart rate							
51-90*	0						
41-50 or 91-110	1	-0.051	0.077	-0.670	0.504	-0.201	0.099
111-130	2	0.132	0.096	1.380	0.168	-0.056	0.320
<41 or >130	3	0.322	0.140	2.290	0.022	0.046	0.597
Systolic BP							
111-219*	0						
101-110	1	0.243	0.108	2.250	0.024	0.032	0.455
91-100	2	0.660	0.144	4.590	0.000	0.378	0.941
<91 or >219	3	0.568	0.163	3.490	0.000	0.249	0.887
Temperature							
36.1-38.0*	0						
35.1-36.0 or 38.1-39.0	1	0.270	0.074	3.640	0.000	0.124	0.415
>39.0	2	0.146	0.132	1.100	0.270	-0.113	0.405
<35.1	3	0.829	0.232	3.570	0.000	0.374	1.283
Consciousness not alert	3	0.411	0.082	5.010	0.000	0.250	0.572
Supplemental oxygen	2	1.197	0.067	17.810	0.000	1.065	1.329
Male sex	1	0.312	0.066	4.710	0.000	0.182	0.442
Age							
<50*	0						
50-65	2	0.737	0.120	6.110	0.000	0.501	0.973
66-80	3	0.828	0.122	6.760	0.000	0.588	1.068
>80	4	0.958	0.129	7.410	0.000	0.705	1.212
Performance status							

Level 1*	0						
Level 2	1	-0.070	0.112	-0.620	0.532	-0.289	0.149
Level 3	2	0.196	0.102	1.920	0.055	-0.004	0.397
Level 4	3	0.413	0.104	3.990	0.000	0.210	0.617
Level 5	4	0.608	0.123	4.960	0.000	0.368	0.848
Renal impairment	1	0.358	0.100	3.590	0.000	0.163	0.553
Respiratory distress	1	0.808	0.153	5.290	0.000	0.509	1.107
Constant		-3.776	0.125	-30.200	0.000	-4.021	-3.531

C-statistic 0.82
*Reference category

Supplementary Table 5: Logistic regression model based on selected categorised predictor variables, excluding respiratory distress and history of renal impairment

Predictor	Score allocated	Coefficient	SE	z	P-value	Upper 95% CI	Lower 95% CI
Respiratory rate							
12-20*	0						
9-11	1	1.500	1.052	1.430	0.154	-0.562	3.562
21-24	2	0.309	0.088	3.500	0.000	0.136	0.483
<9 or >24	3	0.705	0.081	8.710	0.000	0.546	0.863
Oxygen saturation							
>95%*	0						
94-95%	1	0.333	0.089	3.750	0.000	0.159	0.508
92-93%	2	0.637	0.111	5.750	0.000	0.420	0.855
<92%	3	1.116	0.084	13.280	0.000	0.951	1.280
Heart rate							
51-90*	0						
41-50 or 91-110	1	-0.061	0.076	-0.800	0.424	-0.210	0.088
111-130	2	0.122	0.095	1.280	0.200	-0.065	0.309
<41 or >130	3	0.360	0.138	2.600	0.009	0.088	0.631
Systolic BP							
111-219*	0						
101-110	1	0.235	0.108	2.180	0.029	0.024	0.446
91-100	2	0.678	0.143	4.750	0.000	0.398	0.958
<91 or >219	3	0.588	0.162	3.630	0.000	0.271	0.904
Temperature							
36.1-38.0*	0						
35.1-36.0 or 38.1-39.0	1	0.267	0.074	3.620	0.000	0.123	0.412
>39.0	2	0.155	0.131	1.180	0.238	-0.102	0.412
<35.1	3	0.891	0.228	3.910	0.000	0.444	1.339
Consciousness not alert	3	0.406	0.082	4.970	0.000	0.246	0.566
Supplemental oxygen	2	1.224	0.067	18.330	0.000	1.093	1.355
Male sex	1	0.315	0.066	4.770	0.000	0.185	0.444
Age							
<50*	0						
50-65	2	0.752	0.120	6.280	0.000	0.517	0.987
66-80	3	0.850	0.122	6.990	0.000	0.612	1.089
>80	4	0.984	0.128	7.680	0.000	0.733	1.235
Performance status							

Level 1*	0						
Level 2	1	-0.054	0.111	-0.480	0.630	-0.272	0.165
Level 3	2	0.235	0.101	2.330	0.020	0.037	0.434
Level 4	3	0.450	0.103	4.370	0.000	0.248	0.652
Level 5	4	0.655	0.122	5.390	0.000	0.417	0.893
Constant		-3.768	0.124	-30.31	0.000	-4.011	-3.524

C-statistic 0.82
*Reference category

Supplementary Table 6: Sensitivity, specificity, PPV, NPV at each score threshold for predicting the secondary outcome of organ support, validation cohort

Score threshold	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
>0	1.00 (1.00, 1.00)	0.03 (0.03, 0.03)	0.12 (0.12, 0.12)	0.99 (0.98, 1.00)
>1	0.99 (0.99, 1.00)	0.08 (0.08, 0.09)	0.13 (0.12, 0.13)	0.99 (0.99, 0.99)
>2	0.99 (0.98, 0.99)	0.15 (0.15, 0.15)	0.13 (0.13, 0.14)	0.99 (0.99, 0.99)
>3	0.97 (0.97, 0.98)	0.23 (0.22, 0.23)	0.14 (0.14, 0.15)	0.99 (0.98, 0.99)
>4	0.96 (0.95, 0.96)	0.30 (0.29, 0.30)	0.15 (0.15, 0.16)	0.98 (0.98, 0.98)
>5	0.92 (0.91, 0.93)	0.37 (0.37, 0.38)	0.16 (0.16, 0.17)	0.97 (0.97, 0.97)
>6	0.86 (0.85, 0.86)	0.44 (0.44, 0.44)	0.17 (0.17, 0.17)	0.96 (0.96, 0.96)
>7	0.78 (0.77, 0.79)	0.51 (0.50, 0.51)	0.17 (0.17, 0.18)	0.95 (0.94, 0.95)
>8	0.70 (0.69, 0.71)	0.57 (0.57, 0.58)	0.18 (0.17, 0.18)	0.94 (0.93, 0.94)
>9	0.61 (0.60, 0.62)	0.63 (0.63, 0.64)	0.18 (0.18, 0.19)	0.93 (0.92, 0.93)
>10	0.52 (0.51, 0.53)	0.70 (0.69, 0.70)	0.18 (0.18, 0.19)	0.92 (0.91, 0.92)
>11	0.43 (0.42, 0.44)	0.75 (0.74, 0.75)	0.18 (0.18, 0.19)	0.91 (0.91, 0.91)
>12	0.34 (0.33, 0.35)	0.81 (0.80, 0.81)	0.19 (0.18, 0.20)	0.90 (0.90, 0.90)
>13	0.24 (0.23, 0.25)	0.85 (0.85, 0.85)	0.18 (0.17, 0.19)	0.89 (0.89, 0.90)
>14	0.17 (0.16, 0.18)	0.89 (0.89, 0.89)	0.17 (0.16, 0.18)	0.89 (0.89, 0.89)
>15	0.13 (0.12, 0.14)	0.92 (0.92, 0.92)	0.17 (0.16, 0.18)	0.89 (0.89, 0.89)
>16	0.09 (0.08, 0.09)	0.94 (0.94, 0.94)	0.16 (0.15, 0.17)	0.89 (0.88, 0.89)
>17	0.06 (0.05, 0.06)	0.96 (0.96, 0.96)	0.17 (0.15, 0.18)	0.88 (0.88, 0.89)
>18	0.04 (0.03, 0.04)	0.97 (0.97, 0.98)	0.15 (0.14, 0.17)	0.88 (0.88, 0.89)
>19	0.02 (0.01, 0.02)	0.98 (0.98, 0.99)	0.13 (0.11, 0.16)	0.88 (0.88, 0.89)
>20	0.01 (0.01, 0.01)	0.99 (0.99, 0.99)	0.13 (0.10, 0.16)	0.88 (0.88, 0.89)
>21	0.00 (0.00, 0.01)	0.99 (0.99, 0.99)	0.09 (0.06, 0.14)	0.88 (0.88, 0.89)
>22	0.00 (0.00, 0.00)	1.00 (1.00, 1.00)	0.08 (0.04, 0.15)	0.88 (0.88, 0.89)
>23	0.00 (0.00, 0.00)	1.00 (1.00, 1.00)	0.07 (0.02, 0.17)	0.88 (0.88, 0.89)
>24	0.00 (0.00, 0.00)	1.00 (1.00, 1.00)	0.00 (0.00, 0.18)	0.88 (0.88, 0.89)

Supplementary Table 7: Sensitivity, specificity, PPV, NPV at each score threshold for predicting the secondary outcome of death without organ support, validation cohort

Score threshold	Sensitivity (95% CI)	Specificity (95% CI)	Positive predictive value (95% CI)	Negative predictive value (95% CI)
>0	1.00 (1.00, 1.00)	0.03 (0.03, 0.03)	0.13 (0.13, 0.13)	1.00 (0.99, 1.00)
>1	1.00 (1.00, 1.00)	0.09 (0.08, 0.09)	0.13 (0.13, 0.14)	1.00 (1.00, 1.00)
>2	1.00 (1.00, 1.00)	0.15 (0.15, 0.16)	0.14 (0.14, 0.15)	1.00 (1.00, 1.00)
>3	1.00 (0.99, 1.00)	0.23 (0.23, 0.23)	0.16 (0.15, 0.16)	1.00 (1.00, 1.00)
>4	0.99 (0.99, 1.00)	0.30 (0.30, 0.31)	0.17 (0.17, 0.17)	1.00 (1.00, 1.00)
>5	0.98 (0.97, 0.98)	0.38 (0.38, 0.39)	0.18 (0.18, 0.19)	0.99 (0.99, 0.99)
>6	0.96 (0.96, 0.97)	0.46 (0.45, 0.46)	0.20 (0.20, 0.20)	0.99 (0.99, 0.99)
>7	0.93 (0.93, 0.94)	0.53 (0.53, 0.54)	0.22 (0.22, 0.22)	0.98 (0.98, 0.98)
>8	0.89 (0.88, 0.90)	0.60 (0.60, 0.61)	0.24 (0.24, 0.25)	0.97 (0.97, 0.98)
>9	0.84 (0.83, 0.85)	0.67 (0.66, 0.67)	0.26 (0.26, 0.27)	0.97 (0.96, 0.97)
>10	0.77 (0.76, 0.78)	0.73 (0.73, 0.74)	0.29 (0.28, 0.30)	0.96 (0.96, 0.96)
>11	0.70 (0.69, 0.71)	0.79 (0.78, 0.79)	0.32 (0.31, 0.33)	0.95 (0.95, 0.95)
>12	0.59 (0.58, 0.60)	0.84 (0.84, 0.85)	0.35 (0.34, 0.36)	0.94 (0.93, 0.94)
>13	0.49 (0.48, 0.51)	0.89 (0.89, 0.89)	0.39 (0.38, 0.39)	0.93 (0.92, 0.93)
>14	0.40 (0.39, 0.41)	0.92 (0.92, 0.92)	0.43 (0.41, 0.44)	0.92 (0.91, 0.92)
>15	0.32 (0.31, 0.33)	0.95 (0.94, 0.95)	0.46 (0.44, 0.47)	0.91 (0.91, 0.91)
>16	0.25 (0.24, 0.26)	0.96 (0.96, 0.97)	0.50 (0.48, 0.51)	0.90 (0.90, 0.90)
>17	0.17 (0.17, 0.18)	0.98 (0.98, 0.98)	0.52 (0.50, 0.54)	0.89 (0.89, 0.90)
>18	0.11 (0.11, 0.12)	0.99 (0.98, 0.99)	0.52 (0.50, 0.54)	0.89 (0.88, 0.89)
>19	0.07 (0.07, 0.08)	0.99 (0.99, 0.99)	0.61 (0.57, 0.64)	0.88 (0.88, 0.89)
>20	0.05 (0.05, 0.05)	1.00 (1.00, 1.00)	0.64 (0.60, 0.68)	0.88 (0.88, 0.88)
>21	0.03 (0.03, 0.04)	1.00 (1.00, 1.00)	0.72 (0.66, 0.76)	0.88 (0.88, 0.88)
>22	0.02 (0.01, 0.02)	1.00 (1.00, 1.00)	0.76 (0.68, 0.83)	0.88 (0.88, 0.88)
>23	0.01 (0.01, 0.01)	1.00 (1.00, 1.00)	0.80 (0.68, 0.88)	0.88 (0.87, 0.88)
>24	0.01 (0.00, 0.01)	1.00 (1.00, 1.00)	0.86 (0.66, 0.96)	0.88 (0.87, 0.88)

Supplementary Figure 1: Calibration plot for unrestricted LASSO model performance, validation cohort
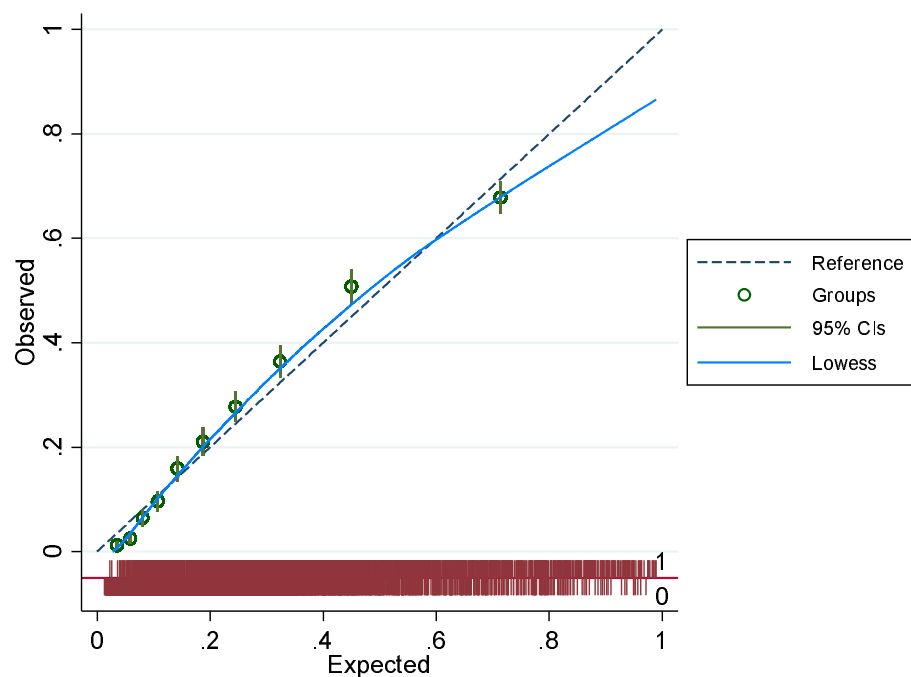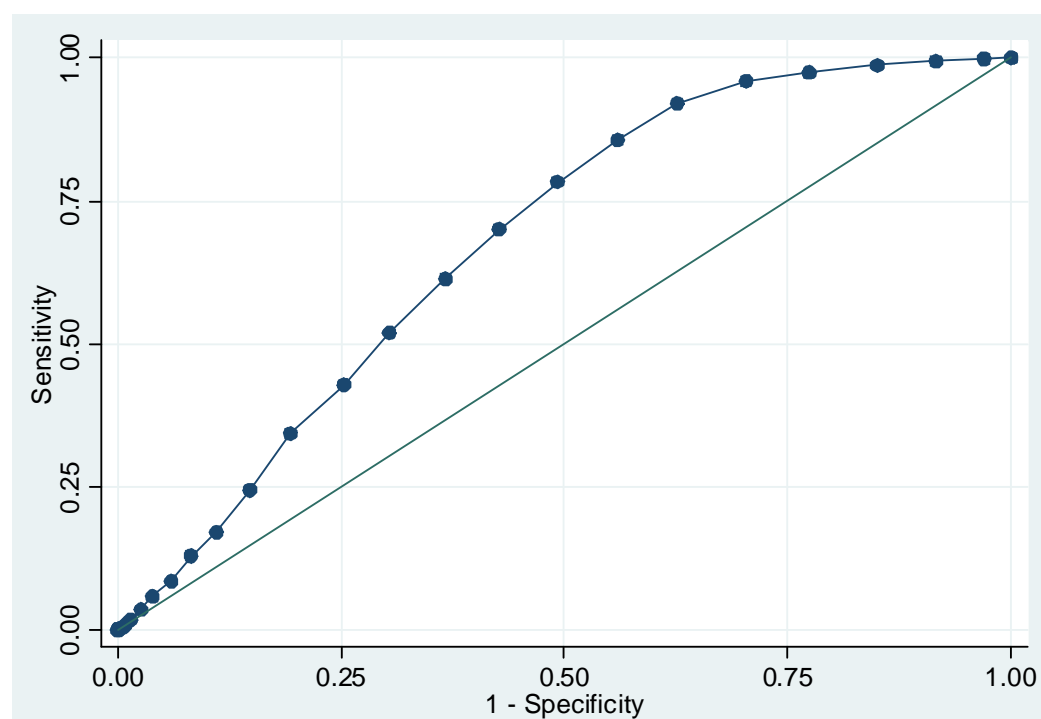

C-statistic 0.82 (95% CI 0.81, 0.83)

Supplementary Figure 2: Calibration plot for restricted LASSO model performance, validation cohort
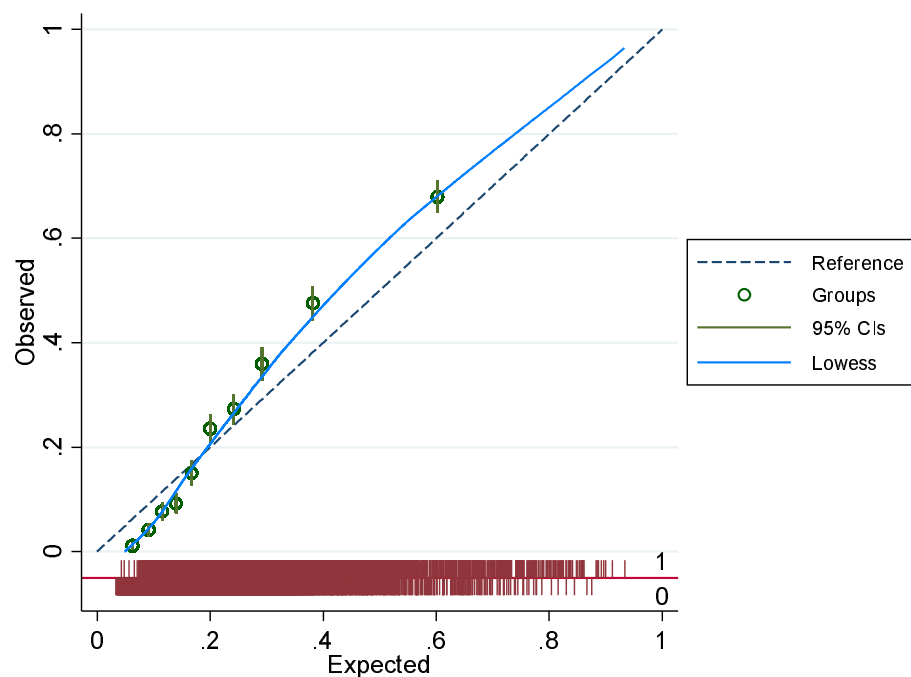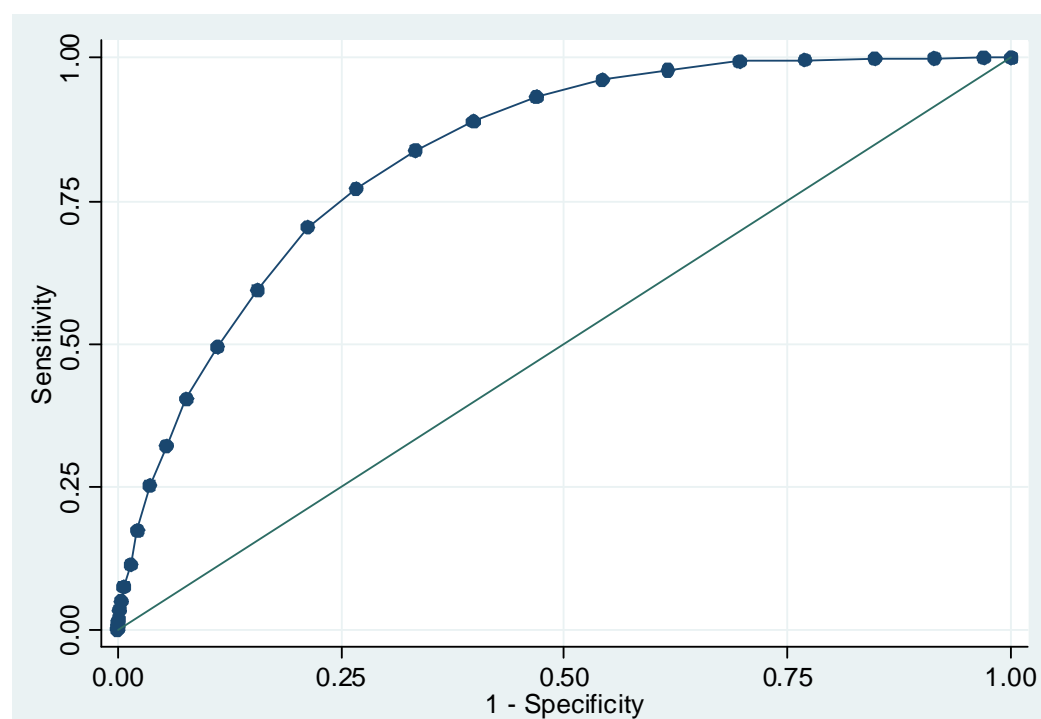

C-statistic 0.81 (95% CI 0.80, 0.82)

Supplementary Figure 3: ROC curve for the tool predicting the secondary outcome of organ support, validation cohort

C-statistic 0.68 (95% CI 0.67, 0.69)

Supplementary Figure 4: ROC curve for tool predicting the secondary outcome of death without organ support, validation cohort

C-statistic 0.83 (95% CI 0.82, 0.84)