



This is a repository copy of *Prognostic accuracy of triage tools for adults with suspected COVID-19 in a pre-hospital setting : an observational cohort study.*

White Rose Research Online URL for this paper:  
<https://eprints.whiterose.ac.uk/176829/>

Version: Submitted Version

---

**Article:**

Marincowitz, C. [orcid.org/0000-0003-3043-7564](https://orcid.org/0000-0003-3043-7564), Sutton, L. [orcid.org/0000-0002-5043-2384](https://orcid.org/0000-0002-5043-2384), Stone, T. [orcid.org/0000-0002-0167-3800](https://orcid.org/0000-0002-0167-3800) et al. (12 more authors) (Submitted: 2021)  
Prognostic accuracy of triage tools for adults with suspected COVID-19 in a pre-hospital setting : an observational cohort study. medRxiv. (Submitted)

<https://doi.org/10.1101/2021.07.27.21261031>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

**Prognostic accuracy of triage tools for adults with suspected COVID-19 in a pre-hospital setting: an observational cohort study.**

1. Carl Marincowitz<sup>1</sup> NIHR Clinical Lecturer Emergency Medicine
2. Laura Sutton<sup>2</sup> Research Associate
3. Tony Stone<sup>1</sup> Senior Data Manager
4. Richard Pilbery<sup>3</sup> Research paramedic
5. Richard Campbell<sup>1</sup> Data Specialist
6. Benjamin Thomas<sup>2</sup> Study Manager
7. Janette Turner<sup>1</sup> Honorary Senior Research Fellow
8. Peter A. Bath<sup>1,4</sup> Professor of Health Informatics and Health Data Analytics
9. Fiona Bell<sup>3</sup> Head of Research
10. Katie Biggs<sup>2</sup> Assistant Director
11. Madina Hasan<sup>1</sup> Research Associate
12. Frank Hopfgartner<sup>4</sup> Senior Lecturer in Data Science
13. Suvodeep Mazumdar<sup>4</sup> Lecturer in Data Analytics
14. Jennifer Petrie<sup>2</sup> Trial Manager
15. Steve Goodacre<sup>1</sup> Professor of Emergency Medicine

1. Centre for Urgent and Emergency Care Research (CURE), Health Services Research School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK, Fax: +44 (0)114 222 0749 Tel: (+44) (0)114 222 4345, Email: [C.Marincowitz@Sheffield.ac.uk](mailto:C.Marincowitz@Sheffield.ac.uk)
2. Clinical Trials Research Unit (CTRU), Health Services Research School of Health and Related Research, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA, UK, Fax: +44 (0)114 222 0749 Tel: (+44) (0)114 222 4345,
3. Yorkshire Ambulance Service NHS Trust, Trust HQ, Springhill 2, Brindley Way, Wakefield, United Kingdom, WF2 0XQ
4. Information School, University of Sheffield, Regent Court, 211 Portobello St, Sheffield S1 4DP, UK

## **Abstract**

### *Study Objective*

Tools proposed to triage patient acuity in COVID-19 infection have only been validated in hospital populations. We estimated the accuracy of five risk-stratification tools recommended to predict severe illness and compare accuracy to existing clinical decision-making in a pre-hospital setting.

### *Methods*

An observational cohort study using linked ambulance service data for patients attended by EMS crews in the Yorkshire and Humber region of England between 18th March 2020 and 29th June 2020 was conducted to assess performance of the PRIEST tool, NEWS2, the WHO algorithm, CRB-65 and PMEWS in patients with suspected COVID-19 infection. The primary outcome was death or need for organ support.

### *Results*

Of 7549 patients in our cohort, 17.6% (95% CI:16.8% to 18.5%) experienced the primary outcome. The NEWS2, PMEWS, PRIEST tool and WHO algorithm identified patients at risk of adverse outcomes with a high sensitivity (>0.95) and specificity ranging from 0.3 (NEWS2) to 0.41 (PRIEST tool). The high sensitivity of NEWS2 and PMEWS was achieved by using lower thresholds than previously recommended. On index assessment, 65% of patients were transported to hospital and EMS decision to transfer patients achieved a sensitivity of 0.84 (95% CI 0.83 to 0.85) and specificity of 0.39 (95% CI 0.39 to 0.40).

### *Conclusion*

Use of NEWS2, PMEWS, PRIEST tool and WHO algorithm could improve sensitivity of EMS triage of patients with suspected COVID-19 infection. Use of the PRIEST tool would improve sensitivity of triage without increasing the number of patients conveyed to hospital.

## Background

Emergency Medical Service (EMS) and other urgent and emergency care practitioners assessing patients with suspected COVID-19 infection in the community, must rapidly determine whether patients need treatment in hospital or can safely remain at home. The overall risk of mortality in patients with confirmed infection is around 1% and if conveyance is too liberal, hospitals could be overwhelmed by patients who require no specific treatment.<sup>1</sup> However, failing to identify a patient at risk of serious deterioration could lead to avoidable harm.<sup>2</sup>

Prognostic research has almost exclusively been conducted in hospital settings and current national and international guidelines for risk-stratification of patients with suspected COVID-19 in the community are consensus based.<sup>1, 3-5</sup> Clinical acuity scores, such as the UK Royal College of Physicians Early Warning Score, version 2 (NEWS2), have been suggested in some guidelines as a way to risk-stratify patients with suspected COVID-19 infection in the community.<sup>6</sup> The WHO decision-making algorithm for respiratory infection and CRB-65 are used to risk-stratify patients with bacterial pneumonia and PMEWS for use in patients with influenza.<sup>7-9</sup> However, the accuracy of these risk-stratification tools has only been validated in hospitalised or non-COVID populations.

NEWS2 has shown good prediction of adverse outcome in patients attending the Emergency Department (ED) with suspected COVID-19.<sup>7</sup> The PRIEST tool was derived by adding age, sex, and performance status to NEWS2, and validation showed improved prediction compared to NEWS2 alone.<sup>10, 11</sup> Validation of the PRIEST tool, NEWS2 and other clinical risk-stratification tools recommended for use in hospital in a community setting,<sup>7, 10, 12-14</sup> could identify the most accurate means to triage need for hospitalisation, thereby reducing unnecessary hospital attendances and improve the identification of those most at risk of serious adverse outcomes.

Our study aimed to:

- 1) Estimate the accuracy of risk-stratification tools recommended to predict severe illness in adults with suspected COVID-19 infection in a pre-hospital setting.
- 2) Compare the accuracy of risk-stratification tools to existing clinical decision-making around transport to hospital.

## Methods

### *Study Design*

This observational cohort study used linked routinely collected EMS data to assess the accuracy in a community setting of five clinical risk-stratification tools (PRIEST tool, NEWS2, WHO algorithm, CRB-

65 and PMEWS) recommended for use in hospitalised patients with COVID-19 or similar respiratory infections (triage tools shown in Supplementary Material 1).<sup>7, 10, 12-14</sup>

### *Setting*

Patients with suspected COVID-19 infection attended by emergency medical services provided by Yorkshire Ambulance Service NHS Trust (YAS). Emergency services provided by YAS covers a region in the north of England of approximately 6,000 square miles and with a population of 5.3 million.

### *Data Sources and linkage*

EMS providers complete an electronic patient report form (ePRF) each time they attend an emergency call, which records presenting patient characteristics and clinical care in a standardised manner. YAS provided a dataset of ePRF data for all EMS responses between the 18th March 2020 and 29th June 2020 where the attending ambulance staff recorded a clinical impression of suspected or confirmed COVID-19 infection. The dataset consisted of patient identifiers, demographic data, measured physiological parameters, other available clinical information and the outcome of the assessment (including whether the patient was conveyed to hospital).

Health and social care data relating to the population in England within the UK National Health Service (NHS) is managed by NHS Digital. We provided patient identifiers to NHS Digital to trace patients in our cohort and supply additional individual level demographic, co-morbidity and outcome data. NHS Digital identified records in their collections belonging to patients in our cohort, and provided data on patient demographics, limited COVID-related general practice (GP) records, emergency department attendances, hospital inpatient admissions, critical care periods, and death registrations from the UK Office of National Statistics.

YAS and NHS Digital removed records where patients indicated that they did not wish their data to be used for research purposes, via the NHS data opt-out service.<sup>15</sup> The study team also excluded patients who had opted out of any part of the PRIEST study and those with inconsistent records (e.g. multiple deaths recorded or death before latest activity). Patient identifiers across all datasets were replaced with a consistent pseudo-identifier to enable the identification and linkage of records belonging to the same patient across all datasets but without revealing any patient's identity.

### *Inclusion Criteria*

Our final cohort consisted of all adult (aged 16 years and over) patients at time of first (index) EMS attendance between 18<sup>th</sup> March and 29<sup>th</sup> June 2020, in which the attending ambulance staff

recorded a clinical impression of suspected or confirmed COVID-19 infection, and who were successfully traced by NHS Digital.

### *Outcome*

The primary outcome was death, renal, respiratory, or cardiovascular organ support (identified from death registration and critical care data) at 30 days from index attendance.

The secondary outcome was death up to 30 days from index contact.

### *Patient Characteristics*

Physiological parameters were extracted from the first (primary) set of clinical observations recorded by the ambulance crew. Consistent with methods used to estimate the Charlson comorbidity index from the available routine data, comorbidities were included if recorded 12 months before the index EMS attendance.<sup>16,17</sup> In a similar way, only immunosuppressant drug prescriptions documented in GP records within 30 days before the index attendance, contributed to the immunosuppression co-morbidity variable. Pregnancy status was based on GP records recorded in the previous 9 months. Frailty in patients older than 65 years was derived from the latest recorded Clinical Frailty Scale (CFS) score (if recorded) in the electronic GP records prior to index attendance.<sup>18</sup> Patients under the age of 65 years were not given a CFS score since it is not validated in this age group. However, if a CFS score was required to calculate a triage tool and the patient was under the age of 65, it was assumed to be 1. Performance status was estimated from the CFS.

### *Analysis*

We retrospectively applied the 5 triage tools to our cohort to assess their accuracy for the primary and secondary outcomes.<sup>7, 10, 12-14</sup> Supplementary Material 1 provides details of scoring and handling missing data for the triage tools. For each tool we plotted the receiver operating characteristic (ROC) curve and calculated the area under the ROC curve (c-statistic) for discriminating between patients with and without adverse outcome. We calculated sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) at the following pre-specified decision making thresholds based on recommended or usual use: 0 vs 1+ CRB-65; 0-1 vs 2+ NEWS2; 0-2 vs 3+ PMEWS; 0-4 vs 5+ PRIEST; 0 vs 1 WHO score. The NEWS2 and PMEWS thresholds used are lower than previously proposed (0-3 vs 4+ NEWS and 0-3 vs 4+ PMEWS) for triaging patient acuity, and are based on the assessment of their performance in a UK ED population of patients with suspected COVID-19 infection, where higher thresholds gave sub-optimal sensitivity.<sup>19</sup> The WHO algorithm and CRB-65

are positive if any criterion is positive. These tools were compared to the sensitivity, specificity, PPV and NPV of EMS clinicians' decision to transfer patients to hospital. All analyses were based on assessment during the index EMS attendance and completed with SAS v9.4.

### **Ethical Approval**

The North West—Haydock Research Ethics Committee gave a favourable opinion on the PAINTED study on 25<sup>th</sup> June 2012 (reference 12/NW/0303) and on the updated PRIEST study on 23rd March 2020, including the analysis presented here. The Confidentiality Advisory Group of the NHS Health Research Authority granted approval to collect data without patient consent in line with Section 251 of the National Health Service Act 2006. Access to data collected by NHS Digital was recommended for approval by its Independent Group Advising on the Release of Data (IGARD) on 11<sup>th</sup> September 2021 having received additional recommendation for approval for access to GP records from the Profession Advisory Group (PAG) on 19<sup>th</sup> August 2021.

### **Patient Public Involvement**

The Sheffield Emergency Care Forum (SECF) is a public representative group interested in emergency care research.<sup>20</sup> 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 conduct of the study.

### **Results**

All totals presented from NHS Digital derived data sets (sex, number of current medications, comorbidities, clinical frailty scores and outcomes) are rounded to the nearest 5, with small numbers suppressed to comply with NHS Digital data disclosure guidance.

#### *Study population*

Figure 1 and Table 1 summarise study cohort derivation and the characteristics of the 7,549 included individual adult patients. In total, 1,330 patients (17.6%, 95% CI:16.8% to 18.5%) experienced the primary outcome (death or organ support) and 1,065 (14.1%, 95% CI: 13.4% to 14.9%), the secondary outcome (death). Of the 7,549 patients, the decision was made to transport 4,905 (65%) to hospital at index attendance. Of those, 1,120 (22.9%) experienced the primary adverse outcome. Of those not transported to hospital, 210 (7.9%) had an adverse outcome. Within the cohort, 3,925

patients (52%, 95% CI:50.9% to 53.1%) were admitted as inpatients and 2,785 (36.9%, 95% CI: 35.8% to 38%) had a diagnosis of COVID confirmed in hospital (since unrestricted community testing was not available until the 18/05/2020) within 30 days of index EMS attendance.

Table 1. Patient characteristics by outcome

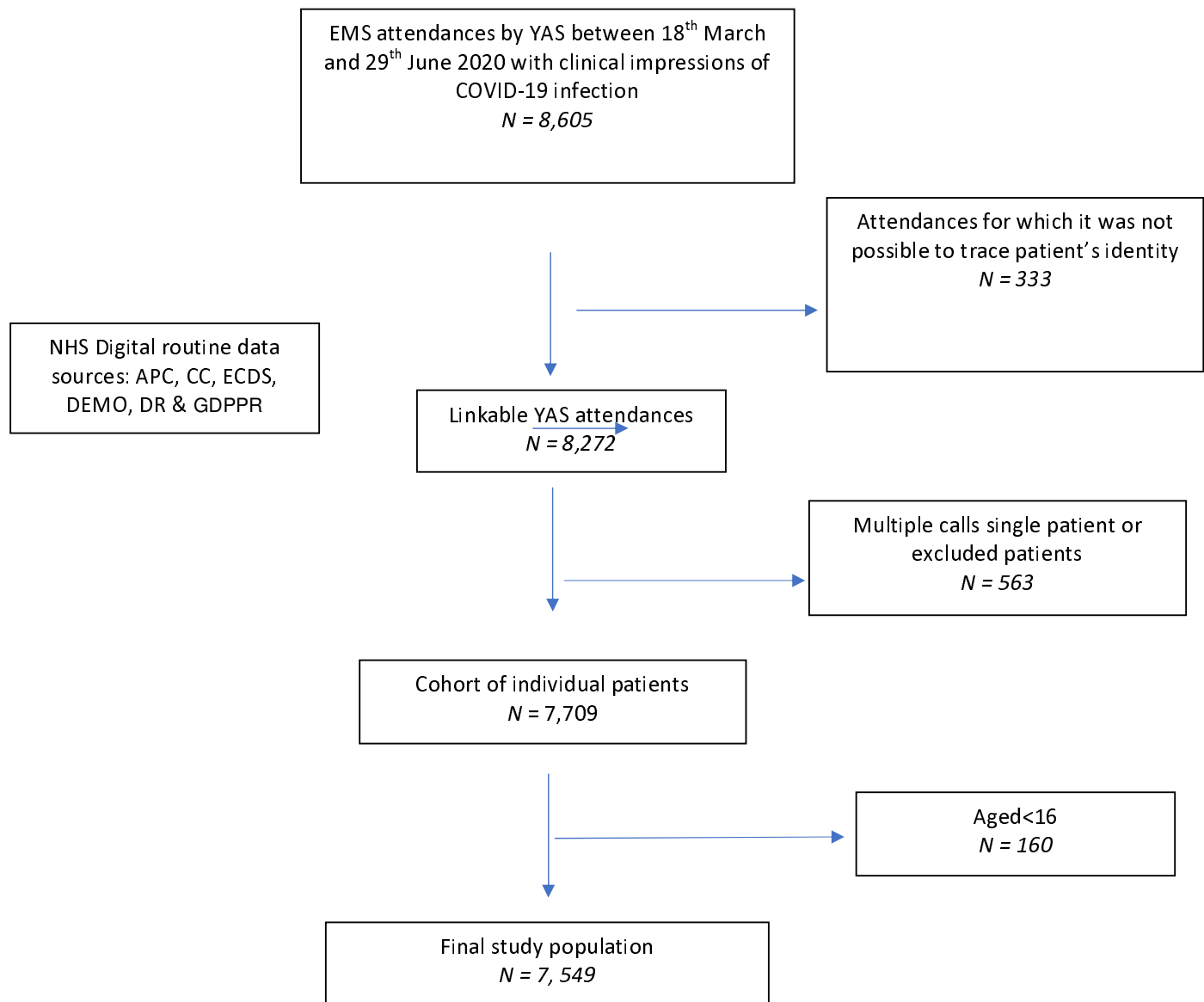
Characteristic	Statistic/level	Adverse outcome	No adverse outcome	Total
	N	1330 (17.6%)*	6220 (82.4%)*	7549
Age (years)*	Mean (SD)	74.5 (15.4)	56.9 (19.4)	60 (20)
	Median (IQR)	78 (65,86)	56 (42,73)	59 (45,77)
	Range	19 to 103	16 to 105	16 to 105
Sex*	Male	760 (57.3%)	2825 (45.4%)	3590 (47.5%)
	Female	570 (42.7%)	3390 (54.6%)	3960 (52.5%)
Number of current medications*	N	1330	6220	7549
	Mean (SD)	4.5 (3.3)	3.2 (3.3)	3.4 (3.3)
	Median (IQR)	4 (2,7)	2 (0,5)	3 (0,6)
	Range	0 to 19	0 to19	0 to19
Comorbidities*	Cardiovascular disease	95 (7%)	290 (4.6%)	380 (5.1%)
	Chronic respiratory disease	375 (28%)	1855 (29.8%)	2230 (29.5%)
	Diabetes	390 (29.2%)	995 (16%)	1380 (18.3%)
	Hypertension	610 (45.8%)	1765 (28.4%)	2375 (31.4%)
	Immunosuppression	280 (21.1%)	930 (15%)	1215 (16.1%)
	Active malignancy	60 (4.6%)	115 (1.9%)	180 (2.3%)
	Renal impairment	55 (4.1%)	125 (2%)	180 (2.4%)
	Stroke	30 (2.3%)	85 (1.4%)	115 (1.5%)
Clinical frailty*	N/A (age <65 years)	330 (47.5%)	3985 (86.4%)	4310 (81.3%)
	Missing	645	1605	2250
	1-3	20 (4.7%)	40 (6.4%)	60 (5.8%)
	4-6	75 (20.5%)	240 (37.7%)	310 (31.4%)
	7-9	270 (74.8%)	350 (55.9%)	620 (62.8%)
AVPU	Missing	13	58	71
	Alert	1002 (76%)	5860 (95.1%)	6862 (91.8%)
	Confusion	125 (9.5%)	188 (3.1%)	313 (4.2%)
	Voice	100 (7.6%)	84 (1.4%)	184 (2.5%)
	Pain	64 (4.9%)	21 (0.3%)	85 (1.1%)
	Unresponsive	27 (2%)	7 (0.1%)	34 (0.5%)
	Missing	13	58	71
Glasgow Coma Scale	N	1297	6085	7382
	Mean (SD)	13.7 (2.4)	14.8 (0.8)	14.6 (1.3)
	Median (IQR)	15 (14,15)	15 (15,15)	15 (15,15)
	Range	3 to 15	3 to 15	3 to 15



Characteristic	Statistic/level	Adverse outcome	No adverse outcome	Total
Diastolic BP (mmHg)	N	1278	6029	7307
	Mean (SD)	76.7 (17.7)	84.5 (15.9)	83.1 (16.5)
	Median (IQR)	76 (65,87)	84 (74,94)	83 (72,93)
	Range	0 to 193	22 to 167	0 to 193
Systolic BP (mmHg)	N	1277	6032	7309
	Mean (SD)	133.2 (25.8)	140.2 (23.2)	139 (23.9)
	Median (IQR)	132 (116,148)	139 (124,153)	138 (123,152)
	Range	65 to 238	33 to 237	33 to 238
Pulse rate (beats/min)	N	1303	6130	7433
	Mean (SD)	100.2 (22.5)	96 (19.5)	96.7 (20.1)
	Median (IQR)	99 (84,115)	94 (82,109)	95 (82,110)
	Range	38 to 194	7 to 190	7 to 194
Respiratory rate (breaths/min)	N	1315	6145	7460
	Mean (SD)	30.1 (10)	23.1 (6.9)	24.4 (8)
	Median (IQR)	28 (22,36)	20 (18,26)	22 (18,28)
	Range	0 to 76	0 to 84	0 to 84
Oxygen saturation	Missing	36	109	145
	>95% on air	142 (11%)	3532 (57.8%)	3674 (49.6%)
	94-95% on air	134 (10.3%)	854 (14%)	988 (13.3%)
	92-93% on air	109 (8.4%)	449 (7.3%)	558 (7.5%)
	<92% on air or O <sub>2</sub> given	910 (70.3%)	1274 (20.9%)	2184 (29.5%)
Blood glucose (mmol/L)	N	982	4021	5003
	Mean (SD)	8.1 (4)	6.9 (3.2)	7.2 (3.4)
	Median (IQR)	6.8 (5.6,9)	6 (5.2,7.3)	6.2 (5.2,7.7)
	Range	0.9 to 35	1.1 to 33.8	0.9 to 35
Temperature (°C)	N	1301	6115	7416
	Mean (SD)	38.1 (1.2)	37.8 (1.1)	37.8 (1.1)
	Median (IQR)	38.2 (37.4,38.9)	37.7 (37,38.5)	37.8 (37,38.6)
	Range	32 to 42	34 to 41.7	32 to 42

\*To comply with NHS digital disclosure guidance totals for these variables are rounded to the nearest 5, which may result in apparent disparities in the overall totals

Figure 1. STROBE flow diagram of study population selection



### *Triage tool performance*

Sensitivity, specificity, positive and negative predictive values for predicting the primary composite outcome using pre-defined score thresholds are provided in Table 2 and the secondary outcome of death in Table 3. Sensitivity and specificity statistics are provided for every score threshold in Supplementary Material 2. The ROC curves for these analyses are shown in Figures 2–3.

Table 2. Triage tool diagnostic accuracy statistics (95% CI) for predicting any adverse outcome

Tool	N*	C-statistic	Threshold	Proportion with score	Sensitivity	Specificity	PPV	NPV
CRB-65	7469	0.79 (0.78, 0.80)	>0	0.54	0.89 (0.88, 0.89)	0.54 (0.53, 0.54)	0.29 (0.29, 0.30)	0.96 (0.95, 0.96)
NEWS2	7433	0.80 (0.78, 0.81)	>1	0.75	0.96 (0.96, 0.96)	0.30 (0.29, 0.30)	0.23 (0.22, 0.23)	0.97 (0.97, 0.97)
PMEWS	7460	0.81 (0.80, 0.83)	>2	0.72	0.98 (0.97, 0.98)	0.34 (0.33, 0.34)	0.24 (0.24, 0.24)	0.99 (0.98, 0.99)
PRIEST	7471	0.83 (0.82, 0.84)	>4	0.66	0.97 (0.97, 0.97)	0.41 (0.40, 0.41)	0.26 (0.25, 0.26)	0.98 (0.98, 0.99)
WHO	7471	0.64 (0.64, 0.65)	>0	0.74	0.98 (0.97, 0.98)	0.31 (0.30, 0.31)	0.23 (0.23, 0.24)	0.98 (0.98, 0.99)

\*Patients with 3 or more missing triage tool parameters were excluded from analysis when estimating performance

The PRIEST tool was robust to the removal of the performance status parameter; when doing so the C-statistic remained at 0.83 (95% CI 0.82 to 0.84). EMS decision to transfer patients to hospital had a sensitivity of 0.84 (95% CI 0.83 to 0.85) and specificity 0.39 (95% CI 0.39 to 0.40) for the primary outcome. The positive predictive value (PPV) was 0.23 (95% CI 0.22 to 0.23) and the negative predictive value (NPV) 0.92 (95% CI 0.92 to 0.92). Hypothetical use of any of the 5 triage tools would have achieved a higher sensitivity than the decision to transfer to hospital by the EMS crews within the cohort, but in the case of NEWS2, the WHO algorithm and PMEWS, this was at a cost of a lower specificity (Table 2). Of the tools assessed at the pre-determined thresholds, CRB-65 achieved the highest specificity but at the cost of sensitivity, and the PRIEST tool achieved a balance between both sensitivity, specificity and C-statistic of 0.83 (95% CI 0.82 to 0.84). The triage tools generally demonstrated better discrimination (except NEWS2) and a higher sensitivity for the secondary outcome, but a lower specificity (Table 3).

Table 3. Triage tool diagnostic accuracy statistics (95% CI) for predicting death within 30 days

Tool	N*	C-statistic	Threshold	Proportion with score	Sensitivity	Specificity	PPV	NPV
CRB-65	7469	0.84 (0.83, 0.85)	>0	0.54	0.95 (0.95, 0.96)	0.53 (0.53, 0.54)	0.25 (0.25, 0.26)	0.99 (0.98, 0.99)
NEWS2	7433	0.78 (0.77, 0.80)	>1	0.75	0.95 (0.95, 0.96)	0.28 (0.28, 0.29)	0.18 (0.18, 0.18)	0.97 (0.97, 0.98)
PMEWS	7460	0.81 (0.80, 0.83)	>2	0.72	0.98 (0.98, 0.98)	0.32 (0.32, 0.33)	0.19 (0.19, 0.20)	0.99 (0.99, 0.99)
PRIEST	7471	0.85 (0.84, 0.86)	>4	0.66	0.98 (0.97, 0.98)	0.39 (0.39, 0.40)	0.21 (0.20, 0.21)	0.99 (0.99, 0.99)
WHO	7471	0.65 (0.64, 0.65)	>0	0.74	0.99 (0.99, 1.00)	0.30 (0.30, 0.30)	0.19 (0.19, 0.19)	1.00 (1.00, 1.00)

\*Patients with 3 or more missing triage tool parameters were excluded from analysis when estimating performance

Figure 2. ROC curves showing triage tool performance for predicting any adverse outcome

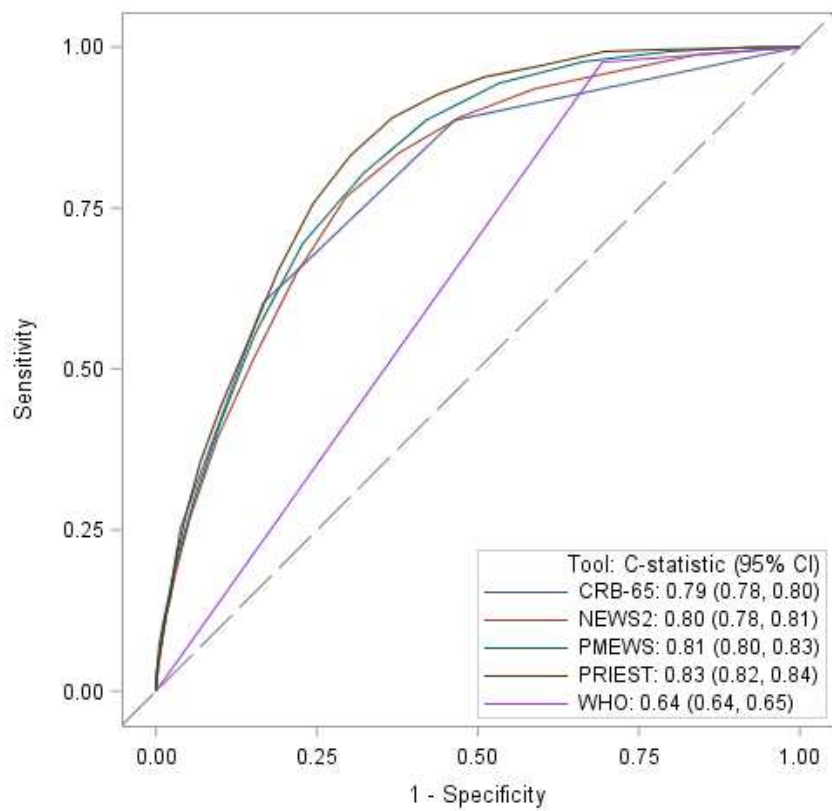
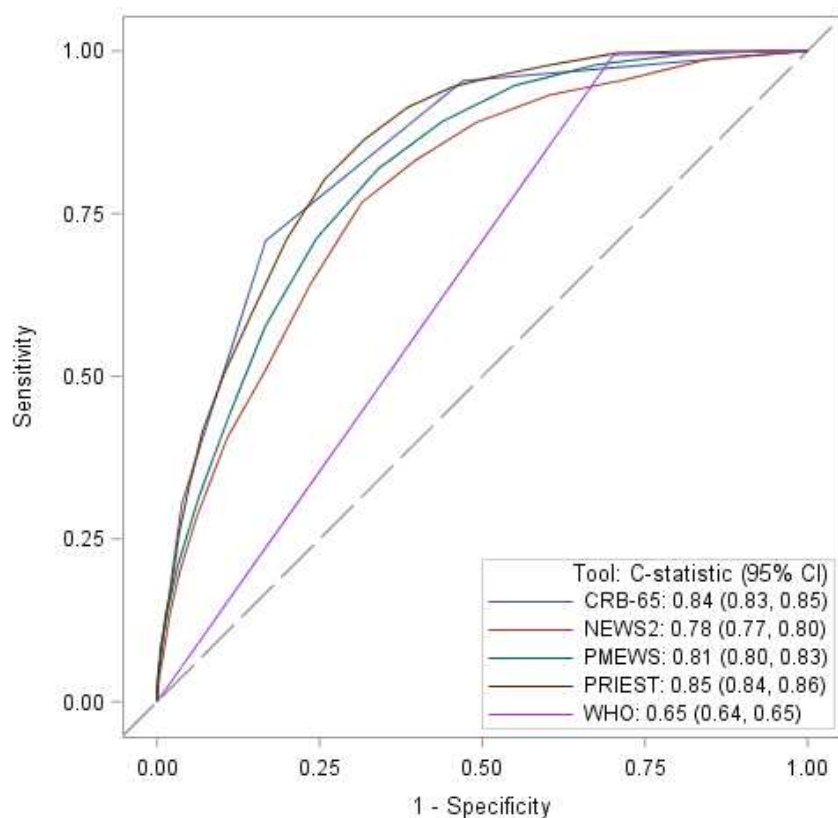


Figure 3. ROC curves showing triage tool performance for predicting death within 30 days



## Discussion

### Summary

The NEWS2, PMEWS, PRIEST tool and WHO algorithm identified patients at risk of adverse outcome with high sensitivity (>0.95) and specificity ranging between 0.3 (NEWS2) and 0.41 (PRIEST tool). They are therefore potentially suitable for use as triage tools to select patients for transfer to hospital. The high sensitivity of NEWS2 and PMEWS was achieved by using lower thresholds (NEWS2; 0–1 vs 2+ and PMEWS; 0–2 vs 3+) than previously recommended, based upon performance in an ED population of patients with suspected COVID-19 infection.<sup>19</sup>

At index attendance, 65% of patients were transported to hospital. Although a useful comparator for triage tool performance, the observed accuracy of EMS decision-making to transfer patients to hospital does not account for clinical best interest decisions not to convey patients to hospital who subsequently deteriorated, or patient wishes not to be conveyed. The sensitivity (0.84, 95% CI 0.83 to 0.85) and specificity (0.39, 95% CI 0.39 to 0.40) achieved by EMS decision making is nonetheless similar to that of tools used to triage undifferentiated patient acuity in the ED.<sup>21</sup>

To be clinically useful to EMS crews, the use of a triage tool would need to improve upon the existing sensitivity of clinical decision making, thereby reducing the risk of not transporting a patient to hospital who subsequently deteriorates, without leading to a disproportionately large increase in hospital conveyance. Use of any of the five triage tools at the pre-specified thresholds would potentially improve upon the sensitivity of existing EMS decision making. However, use of PMEWS, the WHO algorithm or NEWS2 would lead to up to a 10% increase in ED conveyances (Table 2). Use of both CRB65 and the PRIEST score would lead to improvements in sensitivity without sacrificing specificity. CRB65 achieved the highest specificity of any of the tools (0.54 95% CI: 0.53 to 0.54) and its use would reduce the number of patients conveyed to hospital by around 10%. However, patients not conveyed to hospital would have around a 4% risk of subsequently deteriorating. The PRIEST tool achieved a sensitivity of 0.97 (95% CI: 0.97 to 0.97) without increasing the number of patients transported to hospital. Using the PRIEST tool, patients who were not conveyed to hospital would have a 2% risk of subsequent deterioration (compared to an estimated 8% on EMS decision making in this cohort).

#### *Strengths and limitations*

Previous evaluations of triage tool accuracy and prognostic COVID-19 prognostic research in the pre-hospital setting, are limited by only including patients who were subsequently admitted to hospital.<sup>22-25</sup> This is the first evaluation to use a large cohort of patients identified from routinely collected EMS records and linked to nationally collected, patient-level, healthcare data to provide robust outcome data for all patients including those not conveyed to hospital. We had low rates of missing data in the variables used in the triage tools assessed (Table 1). We also assessed the performance of triage tools in a cohort of patients with suspected infection which, in the absence of accurate universally available rapid COVID-19 diagnostic tests, reflects the population which EMS staff must clinically triage. Most existing research either aimed to determine if patients with suspected infection have COVID-19, or to risk stratify patients with confirmed infection in a hospital setting.<sup>26</sup>

Our evaluation of triage tool accuracy is limited to a single ambulance service, albeit one covering a large population across the North of England, so the results may not be generalisable to other healthcare settings. Other ambulance services may serve populations with a different risk-profile, provide different types of EMS response or have different thresholds and guidelines regarding when to convey patients to hospital. The population used is likely to be have similar baseline characteristics to that used to derive and validate the PRIEST score in an ED population, as it was conducted at the same time at hospitals in the region (and elsewhere in the UK).<sup>10</sup> A sensitivity

analysis in which patients recruited to the ED-based PRIEST study were removed from analysis, did not affect estimates of triage tool performance. The PRIEST tool may perform less well if applied to a different, especially non-UK, health care setting.

We assumed that if co-morbidities were not recorded in routine data within the previous 12-months of the index event, they were not present. Our cohort is based on the clinical impression of likely COVID infection as determined by EMS crews. This is partly determined by prevalence of COVID-19 infection which varied during the study period however YAS guidance stated possible COVID-19 infection should be considered in all patients with shortness of breath, cough or fever and in patients with a history of close contact with someone with these symptoms.

### *Implications*

Clinical tools should be used in conjunction with clinical decision making when determining whether a patient needs to be conveyed to hospital by EMS crews. As previously highlighted, there may be good clinical reasons why patients who subsequently deteriorated were not conveyed to hospital in this cohort. Within this limitation, our study provides evidence that use of existing clinical triage tools may improve clinical decision making in a prehospital setting where the prevalence of serious adverse outcomes is similar to the ED.

In health care contexts where minimising risk of adverse outcomes in those not conveyed to hospital is the priority, use of PMEWS or the WHO criteria may be recommended, as they appear to optimise sensitivity. Use of the COVID-specific PRIEST tool would achieve almost the same gains in sensitivity (0.97 versus 0.98) without leading to a corresponding increase in patients being unnecessarily conveyed to hospital. The use of CRB65 would maximise specificity over gains in sensitivity, with a 4% risk of adverse outcomes in patients left to self-care in the community. This may be appropriate in resource constrained health care contexts and, as oxygen saturations do not form part of the assessment tool, it can be practically applied to a large range of health care settings.

Further research assessing triage tool performance alongside clinical judgement in the prehospital setting would be helpful to determine whether triage tools would improve accuracy of decisions to transfer patients to hospital in practice. Given the high prevalence of adverse outcomes in this cohort, the findings may not be applicable to other lower risk community settings (e.g. patients being assessed by general practitioners) and therefore similar research is needed for these populations.

### Conclusion

The NEWS2, PMEWS, PRIEST tool and WHO algorithm achieved high estimated sensitivities (>0.95) with respect to death or organ support, and specificities ranging between 0.3 (NEWS2) and 0.41 (PRIEST tool). EMS decision to transfer patients to hospital achieved a sensitivity of 0.84 (95% CI 0.83 to 0.85). Although, there may be good clinical reasons why patients who deteriorated were not conveyed to hospital, use of these triage tools would potentially improve EMS triage of patients with suspected COVID-19 infection. Use of the PRIEST tool would lead to significant gains in sensitivity without increasing the number of patients conveyed to hospital.

1. Bentivegna M, Hulme C, Ebell MH. Primary Care Relevant Risk Factors for Adverse Outcomes in Patients With COVID-19 Infection: A Systematic Review. *The Journal of the American Board of Family Medicine*. 2021;34:S113.
2. Gandhi RT, Lynch JB, del Rio C. Mild or Moderate Covid-19. *New England Journal of Medicine*. 2020;383:1757-1766.
3. <https://www.who.int/publications/i/item/WHO-2019-nCoV-clinical-2021-1>.
4. <https://www.nice.org.uk/guidance/ng191>.
5. <https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html>.
6. Greenhalgh T, Thompson P, Weiringa S, et al. What items should be included in an early warning score for remote assessment of suspected COVID-19? qualitative and Delphi study. *BMJ Open*. 2020;10:e042626.
7. "Clinical care of severe acute respiratory infections – Tool kit," 11 April 2020. [Online]. Available: <https://www.who.int/publications-detail/clinical-care-of-severe-acute-respiratory-infections-tool-kit>. In: WHO, ed.
8. <https://bestpractice.bmj.com/topics/en-gb/3000108/management-recommendations>.
9. Goodacre S, Challen K, Wilson R, Campbell M. Evaluation of triage methods used to select patients with suspected pandemic influenza for hospital admission: cohort study. *Health Technol Assess*. 2010;14:173-236.
10. Goodacre S, Thomas B, Sutton L, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: The PRIEST observational cohort study. *PLoS One*. 2021;16:e0245840.
11. Thomas B, Goodacre S, Lee E, et al. Prognostic accuracy of emergency department triage tools for adults with suspected COVID-19: The PRIEST observational cohort study. *medRxiv*. 2020:2020.2009.2002.20185892.
12. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS. Updated report of a working party. London: RCP. In: Physicians RCo, ed2017.
13. Lim WS, van der Eerden MM, Laing R, et al. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax*. 2003;58:377-382.
14. 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.
15. <https://digital.nhs.uk/services/national-data-opt-out>.
16. Campbell MJ, Jacques RM, Fotheringham J, Maheswaran R, Nicholl J. Developing a summary hospital mortality index: retrospective analysis in English hospitals over five years. *BMJ*. 2012;344:e1001.



17. Street A, Maynou L, Gilbert T, Stone T, Mason S, Conroy S. The use of linked routine data to optimise calculation of the Hospital Frailty Risk Score on the basis of previous hospital admissions: a retrospective observational cohort study. *The Lancet Healthy Longevity*. 2021;2:e154-e162.
18. Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. *CMAJ*. 2005;173:489-495.
19. Thomas B, Goodacre S, Lee E, et al. Prognostic accuracy of emergency department triage tools for adults with suspected COVID-19: the PRIEST observational cohort study. *Emergency Medicine Journal*. 2021:emermed-2020-210783.
20. Hirst E, Irving A, Goodacre S. Patient and public involvement in emergency care research. *Emergency Medicine Journal*. 2016;33:665-670.
21. Zachariasse JM, van der Hagen V, Seiger N, Mackway-Jones K, van Veen M, Moll HA. Performance of triage systems in emergency care: a systematic review and meta-analysis. *BMJ Open*. 2019;9:e026471.
22. Lancet EA, Gonzalez D, Alexandrou NA, et al. Prehospital hypoxemia, measured by pulse oximetry, predicts hospital outcomes during the New York City COVID-19 pandemic. *Journal of the American College of Emergency Physicians Open*. 2021;2:e12407.
23. Dillon K, Hook C, Coupland Z, Avery P, Taylor H, Lockyer A. Pre-hospital lowest recorded oxygen saturation independently predicts death in patients with COVID-19. *Br Paramed J*. 2020;5:59-65.
24. Saberian P, Tavakoli N, Hasani-Sharamin P, Modabber M, Jamshididana M, Baratloo A. Accuracy of the pre-hospital triage tools (qSOFA, NEWS, and PRESEP) in predicting probable COVID-19 patients' outcomes transferred by Emergency Medical Services. *babol-caspjim*. 2020;11:536-543.
25. Inada-Kim M, Chmiel FP, Boniface MJ, Pocock H, Black JJM, Deakin CD. Validation of home oxygen saturations as a marker of clinical deterioration in patients with suspected COVID-19. *medRxiv*. 2020:2020.2011.2006.20225938.
26. 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.