RESEARCH ARTICLE

Evaluation of a hypothetical decision-support tool for intensive care triage of patients with coronavirus disease 2019 (COVID-19) [version 1; peer review: awaiting peer review]

Emily Simon Thomas1, Bryony Peiris1, Leon Di Stefano2, Matthew J. Rowland1,3, Dominic Wilkinson1,4,5

1John Radcliffe Hospital, Oxford, UK
2Johns Hopkins Bloomberg School of Public Health, Baltimore, USA
3Kadoorie Centre for Critical Care Research, Nuffield Department of Clinical Neurosciences, Oxford University, Oxford, UK
4Oxford Uehiro Centre for Practical Ethics, Oxford University, Oxford, UK
5Murdoch Children’s Research Institute, Melbourne, Australia

Abstract

Background: At the start of the coronavirus disease 2019 (COVID-19) pandemic there was widespread concern about potentially overwhelming demand for intensive care and the need for intensive care unit (ICU) triage. In March 2020, a draft United Kingdom (UK) guideline proposed a decision-support tool (DST). We sought to evaluate the accuracy of the tool in patients with COVID-19.

Methods: We retrospectively identified patients in two groups: referred and not referred to intensive care in a single UK national health service (NHS) trust in April 2020. Age, Clinical Frailty Scale score (CFS), and co-morbidities were collected from patients’ records and recorded, along with ceilings of treatment and outcome. We compared the DST, CFS, and age alone as predictors of mortality and treatment ceiling decisions.

Results: In total, 151 patients were included in the analysis, with 75 in the ICU and 76 in the non-ICU-reviewed groups. Age, clinical frailty and DST score were each associated with increased mortality and higher likelihood of treatment limitation (p-values all <.001). A DST cut-off score of >8 had 65% (95% confidence interval (CI) 51%-79%) sensitivity and 63% (95% CI 54%-72%) specificity for predicting mortality. It had a sensitivity of 80% (70%-88%) and specificity of 96% (95% CI 90%-100%) for predicting treatment limitation. The DST was more discriminative than age alone (p<0.001), and potentially more discriminative than CFS (p=0.08) for predicting treatment ceiling decisions.

Conclusions: During the first wave of the COVID-19 pandemic, in a hospital without severe resource limitations, a hypothetical decision support tool was limited in its predictive value for mortality, but
appeared to be sensitive and specific for predicting treatment limitation.

**Keywords**
COVID-19, ethics, intensive care, triage, withholding treatment

This article is included in the Coronavirus (COVID-19) collection.

**Corresponding authors:** Emily Simon Thomas (esimonthomas@hotmail.com), Dominic Wilkinson (dominic.wilkinson@philosophy.ox.ac.uk)

**Author roles:** Simon Thomas E: Data Curation, Formal Analysis, Methodology, Project Administration, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; Peiris B: Data Curation, Formal Analysis, Investigation, Methodology, Validation, Writing – Review & Editing; Di Stefano L: Formal Analysis, Methodology, Resources, Software, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing; Rowland MJ: Formal Analysis, Investigation, Methodology, Resources, Validation, Writing – Review & Editing; Wilkinson D: Conceptualization, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Supervision, Validation, Visualization, Writing – Original Draft Preparation, Writing – Review & Editing

**Competing interests:** MR is currently employed by Roche on a senior clinical research fellowship.

**Grant information:** This research was supported by the Wellcome Trust [203132]; and by the UKRI/ AHRC funded UK Ethics Accelerator project [AH/V013947/1]. The funders had no role in the preparation of this manuscript or the decision to submit for publication.

**Copyright:** © 2021 Simon Thomas E et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Simon Thomas E, Peiris B, Di Stefano L et al. Evaluation of a hypothetical decision-support tool for intensive care triage of patients with coronavirus disease 2019 (COVID-19) [version 1; peer review: awaiting peer review] Wellcome Open Research 2021, 6:183 https://doi.org/10.12688/wellcomeopenres.16939.1

**First published:** 19 Jul 2021, 6:183 https://doi.org/10.12688/wellcomeopenres.16939.1
Introduction

In the first phase of the coronavirus disease 2019 (COVID-19) pandemic, in March 2020, there was widespread concern in the United Kingdom (UK) that there would be insufficient intensive care unit (ICU) beds and mechanical ventilators to treat the number of patients presenting with severe COVID-19. Pandemic modelling suggested that rates of infection in the UK would exceed ICU capacity.

Even outside of a pandemic, not all patients with severe illness are admitted to intensive care. There are patients too well to benefit from ICU management, and others for whom the short or long-term benefit is uncertain or likely to be small relative to the burden of treatment. In the event of there being intense pressure on ICU capacity, as was feared to be the case during the first wave of COVID-19, there could be a role for a system to support allocation of these resources in a fair and clinically effective way.

The National Institute for Health and Care Excellence (NICE) published a rapid clinical guideline on critical care for adults in the context of COVID-19 on 20th March, 2020. It recommended that all admissions be assessed for frailty, using the Clinical Frailty Scale (CFS). It suggested that patients who are less frail (“for example a CFS score of less than 5”) would likely benefit from intensive care while for more frail patients (e.g. CFS 5 or above) there was uncertainty about benefit.

A draft UK national pandemic allocation guideline, developed by the UK Moral and Ethical Advisory Group (MEAG) in conjunction with the Intensive Care Society in late March 2020, proposed a scoring system incorporating frailty, severity and co-morbidities (Figure 1). This suggested benefit of intubation and ventilation for patients with a Decision Support Tool (DST) score of eight or below, while for patients with a DST of >8 it would be appropriate to limit treatment (potentially including continuous positive airway pressure (CPAP)/non-invasive ventilation). The MEAG guideline and tool was apparently rejected by UK health officials, and no official NHS guidance was produced.

There were two potential justifications for thresholds proposed in the NICE and MEAG guidance. First, that high scores would be predictive of little benefit from treatment in patients with COVID-19 (i.e. low survival). This would then provide a basis for rationing treatment in the event of a critical shortage in the availability of intensive care. Secondly, such scores...
might correlate with the decisions about the appropriateness of intensive care ordinarily made by ICU specialists. The scoring system would therefore allow rapid and consistent decisions by other members of the healthcare team. However, at the time that these were proposed, there were no available data on how well the CFS or DST thresholds would predict mortality from COVID-19, nor whether they would reliably predict decisions about ceilings of treatment.

The aim of this study was to evaluate the proposed MEAG decision-support tool (DST) and CFS threshold for predicting mortality and for identifying patients judged clinically appropriate for intensive care admission, in patients hospitalised with COVID-19 in a hospital without severe resource limitations during the first wave of the pandemic.

Methods
We sought to retrospectively identify all adult patients (aged 17 years and above) with COVID-19 referred in our trust for potential ICU admission in April 2020. Oxford University Hospitals NHS Foundation Trust (OUH) includes one tertiary centre, one specialist hospital and one district general hospital; there are a total of 1465 beds with a baseline of 60 intensive care beds. Medical record numbers of patients reviewed by ICU consultants were collected prospectively.

During this period, decisions about admission to intensive care for patients referred to ITU were made by the on-call ITU consultant on the basis of the patient’s wishes, clinical condition and the perceived appropriateness of intensive care admission. There were no policies setting thresholds for admission, and the DST was neither used nor recorded. Patients receiving only non-invasive ventilation were largely treated outside on intensive care.

We retrospectively identified a control group of patients with COVID-19 admitted to OUH but not referred to intensive care. (Decisions about treatment for this cohort were made by the treating consultant/senior registrar on the basis of the patient’s wishes and clinical appropriateness). Because of the larger number of patients, we restricted the control group to the two-week window from 8th April to 21st April 2020 to arrive at a cohort of the same size as the ITU group. Control patients were identified from patients identified on the trust electronic patient record system. Patients who did not have COVID-19 or ‘COVID-syndrome’ (presumed COVID-19 on the basis of clinical and radiographic findings but tested negative on swabbing) or did not have an oxygen requirement were excluded.

The electronic patient records were examined by two independent data collectors (EST and BP). Disagreements were discussed and resolved by consensus or review by a third researcher (DW). Age, clinical frailty score, and specific co-morbidities identified as part of the DST were recorded (cardiac arrest in last 3 years from any cause, any chronic condition causing three or more hospital admissions in the last year or four or more weeks continuous admission for current inpatients, congestive heart failure or chronic lung disease with symptoms at rest or on minimal exertion, hypertension, severe and irreversible neurological condition, chronic liver disease with Child-Pugh score of ≥7, end stage chronic renal failure requiring renal replacement therapy, diabetes mellitus requiring medication, uncontrolled or active malignancy). Age was classified into subgroups as specified in the DST. At the time of the study, clinical frailty was recorded routinely for many admissions with COVID-19 as per NICE guidance. Where a frailty score was not documented, relevant information from the patient’s electronic records, including from the initial clerking and reviews by physiotherapists and occupational therapists, was used to estimate this (based on consensus between data collectors). A patient who was a marathon runner at the time of admission would for example be assigned a score of 1, a patient who was reliant on support for all of their personal care but stable would have been assigned a score of 7, etc, per the NICE CFS.

Admission and review dates, whether the patient had a treatment escalation plan documented, treatment offered, and ceiling of treatment decisions during admission were recorded. Ceilings of treatment were classified into three categories designating the maximum level of support that would be potentially appropriate to provide: (1) intubation and ventilation, (2) non-invasive ventilation including high flow nasal oxygen and CPAP, and (3) administration of oxygen by other means (e.g. nasal cannula). If ceiling of treatment was not clearly documented, patients were excluded for analysis of this (but were still able to be included in mortality figures).

Hospital outcomes were recorded from electronic records. These were whether, as of the end of May 2020, the patient had been discharged, was still an inpatient, or had died.

This project was registered as an audit with the Trust quality improvement team on 17th April 2020. Ethics review was not required.

We compared the DST, CFS, and age alone as predictors of mortality, and treatment ceiling decision (i.e. ceiling of treatment not including intubation/ventilation). We assessed each predictor’s performance using receiver operating characteristic (ROC) curves and the area under the curve (AUC). We also estimated each predictor’s sensitivity and specificity at specific thresholds: for DST score a score of ≤3 vs >3; for the CFS, a score of ≤4 vs. >4; and for age, ≤75 vs. >75 years old.

For statistical analysis we used the statistical programming language R version 4.0.2 and the package pROC\(^1\). We computed sensitivity and specificity values for ROC curves using that package’s \texttt{roc} function. To generate confidence intervals (CIs) for sensitivity and specificity of each predictor, we used the package’s bootstrap intervals\(^1\). To generate confidence
intervals for AUC values, we used pROC’s bootstrap-based method; to test the pairwise differences between AUC values, we used the package’s implementation of De Long’s test\(^4\).

**Results**

In the month of April 2020, 749 patients over the age of 17 years tested positive for COVID-19 in OUH. A further 1085 had suspected COVID-19 without a laboratory test or with a laboratory test reported as inconclusive i.e. based on symptoms, radiological evidence, other blood test results. In April 2020, 70 patients with COVID-19 were treated in intensive care units in OUH, of which 57 were discharged and 13 died (18.6% mortality)\(^5\).

In total, 151 patients were included in this study, with 75 in the ICU and 76 in the non ICU-reviewed groups (Table 1).

Ceiling of care was not clearly documented in 4 cases (2 ICU reviewed, 2 non-reviewed).

Patients who were reviewed by intensive care were younger, less frail, and had lower decision-support tool scores; Mann-Whitney tests produced negligible \(p\)-values (<.001). (Table 1, Figure 2). There were 4 patients reviewed by ICU with a CFS score <5 and 9 with a DST score >8.

**Decisions and outcome**

Among patients referred for potential ICU admission, 57 (78%) had no treatment limitation documented, (27 of this 57 (47%) were intubated), while 16 (22%) had a ceiling of treatment of non-invasive ventilation documented. In the non ICU-reviewed group, 41 (55%) had a ceiling of treatment of nasal canulae or face mask oxygen delivery, 20 (27%) of non-invasive

---

**Table 1. Demographics of the intensive care unit (ICU) reviewed and non-ICU reviewed patient groups.**

<table>
<thead>
<tr>
<th></th>
<th>ICU reviewed (n=75)</th>
<th>Non-reviewed (n=76)</th>
<th>Total (n=151)</th>
<th>Test of no difference between groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age: mean (±SD, range)</td>
<td>59 (±14.02, 23-83)</td>
<td>76 (±13.8, 39-96)</td>
<td>67 (±16.2, 23-96)</td>
<td>( p = 5.2 \times 10^{-10} ) (Mann-Whitney test using age categories)</td>
</tr>
<tr>
<td>Gender</td>
<td>31F:44M</td>
<td>34F:42M</td>
<td>65F:86M</td>
<td>( p = 0.80 ) (( \chi^2 = 0.07 ))</td>
</tr>
<tr>
<td>Clinical Frailty Score: median IQR (Q1-Q3)</td>
<td>2 (1-2)</td>
<td>5 (1-4)</td>
<td>3 (2.5-5.5)</td>
<td>( p = 1.4 \times 10^{-15} ) (Mann-Whitney)</td>
</tr>
<tr>
<td>Decision Support Tool Score: median IQR (Q1-Q3)</td>
<td>5 (3-7)</td>
<td>12 (9-13)</td>
<td>8 (4-12)</td>
<td>( p = 2.6 \times 10^{-16} ) (Mann-Whitney)</td>
</tr>
</tbody>
</table>

---

**Figure 2. Distribution of decision support tool triage scores among intensive care unit (ICU) referred and non-ICU referred patients.** Orange line indicates the suggested Decision Support Tool (DST) threshold. Y-axis shows number of patients.
ventilation (5/20 received NIV), and 13 (18%) of intubation and ventilation, should that need arise. Seven (9%) were treated in high dependency areas, none were intubated.

Of the ICU-reviewed group, at the end of May 2020, 18 had died (24%), 53 (71%) had been discharged, and 4 were still inpatients. In the non-ICU-reviewed group, 25 had died (33%) while 51 (67%) had been discharged.

**Prediction of mortality**
Increased age, clinical frailty and DST score were each associated with increased mortality (separate logistic regressions of mortality on CFS, DST score, and age category index all produced p-values <.001) (Table 2).

In total, 41% (28/68) of patients with a DST score of 9 or higher died, compared with 18% (15/83) of patients with a score of 8 or below (Table 3). A cut-off score of >8 had 65% (95% CI 51% - 79%) sensitivity for predicting mortality, and 63% (95% CI 54% - 72%) specificity (Figure 3).

In comparison, 44% [26/59] of patients with a clinical frailty score of five or higher died, while 18% [17/92] of those with a clinical frailty score below five died. A CFS of ≥5 had a 60% (95% CI 47% - 77%) sensitivity for predicting mortality and 69% (95% CI 61% - 78%) specificity. The area under the curve (AUC) for prediction of mortality was 0.70 for age categories (95% CI 0.61-0.79), 0.68 for the CFS (95% CI 0.58- 0.77) and 0.72 for the DST (95% CI 0.64-0.81). There were no statistically significant pairwise differences between the three metrics as discriminators of mortality (Figure 3–Figure 5).

**Prediction of clinical decision**
Increased age, clinical frailty and DST score were each associated with a higher likelihood of a treatment limitation decision (p-values from separate logistic regressions of treatment decision on CFS, DST score, and age category index all <.001).

Of patients with a DST equal to or below 8, 15% (12/79) had a treatment limitation decision. Overall, 96% (65/68) of patients with a DST of more than 8 had a treatment limitation decision (Table 2). A DST cut-off of more than 8 had a sensitivity of 80% (70%-88%) and a specificity of 96% (95% CI 90%-100%) for predicting treatment limitation.

In comparison, 22% (19/88) of patients with a CFS of <5 had a treatment limitation decision. 98% (58/59) of patients with a CFS of 5 or above had a treatment limitation decision. A CFS value of 5 or above had a sensitivity of 72% (95% CI 60%-81%) and a specificity of 99% (95% CI 96%-100%) for predicting treatment limitation.

The AUC for prediction of a treatment limitation decision was 0.87 (95% CI 0.81-0.92) for age category, 0.91 for CFS (95% CI 0.87- 0.95) and 0.95 (95% CI 0.91- 0.98) for DST (Figure 3).

**Discussion**
In this single centre study, among patients hospitalised with COVID-19 in a UK hospital during the first wave of the pandemic, a hypothetical triage scoring tool (DST) had moderate sensitivity and specificity for predicting in-hospital mortality (65% and 63%). The tool appeared to be highly predictive of clinical decisions about the appropriateness of intubation and ventilation, made in the absence of severe resource shortage. The DST (which includes components of age, comorbidities and clinical frailty) appeared to be more discriminatory than age or clinical frailty alone.

When the study was first conceived, it was unclear whether a national guideline in relation to intensive care triage would be forthcoming. It was felt to be important, if possible, to evaluate the accuracy of such a tool before it was adopted. On 25th October 2020 an official statement was issued by the NHS in response to allegations that had been made that patients who were frail and elderly had been denied adequate care during the first wave of the COVID-19 pandemic. The statement emphasised that the NHS had not run out of critical care beds in this time period, and that intensive care was offered to the patients who were assessed and deemed likely to benefit from it on an individual basis. It acknowledged that a tool had been developed to facilitate triaging of patients, but that this was neither formally launched nor was use of the NHS logo on it authorised.

Previous pandemic plans have focused on illness severity scores, sometimes in combination with exclusion of patients with severe life-limiting comorbidity (e.g. metastatic malignant disease). For example, the Ontario Health Plan for an Influenza epidemic (OHPIP) was recommended for use in a future influenza pandemic. It combines Sequential Organ
Table 3. Mortality and treatment limitation per clinical frailty (CFS) and Decision Support Tool (DST) scores.
ICU=intensive care unit.

<table>
<thead>
<tr>
<th>CFS</th>
<th>ICU (n=75)</th>
<th>Non-ICU (n=76)</th>
<th>Total (n=151)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mortality</td>
<td>treatment limitation</td>
<td>mortality</td>
</tr>
<tr>
<td>1</td>
<td>0/5 (0%)</td>
<td>0/5 (0%)</td>
<td>0/3 (0%)</td>
</tr>
<tr>
<td>2</td>
<td>6/35 (17%)</td>
<td>1/34* (3%)</td>
<td>2/3 (50%)</td>
</tr>
<tr>
<td>3</td>
<td>6/26 (23%)</td>
<td>8/25* (32%)</td>
<td>1/11 (9%)</td>
</tr>
<tr>
<td>4</td>
<td>1/5 (20%)</td>
<td>2/5 (40%)</td>
<td>0/4 (0%)</td>
</tr>
<tr>
<td>5</td>
<td>1/1 (100%)</td>
<td>0/1 (0%)</td>
<td>5/20 (25%)</td>
</tr>
<tr>
<td>6</td>
<td>3/3 (100%)</td>
<td>3/3 (100%)</td>
<td>7/17 (41%)</td>
</tr>
<tr>
<td>7</td>
<td>0/0</td>
<td>0/0</td>
<td>8/16 (50%)</td>
</tr>
<tr>
<td>8</td>
<td>0/0</td>
<td>0/0</td>
<td>0/1 (0%)</td>
</tr>
<tr>
<td>9</td>
<td>0/0</td>
<td>0/0</td>
<td>1/1 (100%)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DST</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0/1 (0%)</td>
<td>0/1 (0%)</td>
<td>0/1 (0%)</td>
<td>0/2 (0%)</td>
<td>0/2 (0%)</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0/5 (0%)</td>
<td>0/5 (0%)</td>
<td>0/0</td>
<td>0/5 (0%)</td>
<td>0/5 (0%)</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0/9 (0%)</td>
<td>0/8* (0%)</td>
<td>0/1 (0%)</td>
<td>0/11 (0%)</td>
<td>0/11 (0%)</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1/12 (8%)</td>
<td>0/12 (0%)</td>
<td>0/3 (0%)</td>
<td>0/2* (0%)</td>
<td>1/13 (8%)</td>
<td>0/12* (0%)</td>
</tr>
<tr>
<td>4</td>
<td>1/10 (10%)</td>
<td>1/10 (10%)</td>
<td>0/1 (0%)</td>
<td>0/1 (0%)</td>
<td>1/12 (7%)</td>
<td>1/12 (7%)</td>
</tr>
<tr>
<td>5</td>
<td>2/10 (20%)</td>
<td>2/10 (18%)</td>
<td>0/4 (0%)</td>
<td>0/4 (0%)</td>
<td>2/14 (12%)</td>
<td>2/14 (12%)</td>
</tr>
<tr>
<td>6</td>
<td>1/5 (20%)</td>
<td>1/5 (20%)</td>
<td>2/3 (67%)</td>
<td>1/3 (33%)</td>
<td>3/8 (38%)</td>
<td>2/8 (25%)</td>
</tr>
<tr>
<td>7</td>
<td>4/7 (57%)</td>
<td>2/7 (29%)</td>
<td>0/1 (0%)</td>
<td>0/0*</td>
<td>4/8 (50%)</td>
<td>2/7* (29%)</td>
</tr>
<tr>
<td>8</td>
<td>4/7 (57%)</td>
<td>3/6* (50%)</td>
<td>0/3 (0%)</td>
<td>2/3 (67%)</td>
<td>4/10 (40%)</td>
<td>5/9* (56%)</td>
</tr>
<tr>
<td>9</td>
<td>1/5 (20%)</td>
<td>3/5 (60%)</td>
<td>0/4 (0%)</td>
<td>4/4 (100%)</td>
<td>1/9 (11%)</td>
<td>7/9 (78%)</td>
</tr>
<tr>
<td>10</td>
<td>3/3 (100%)</td>
<td>3/3 (100%)</td>
<td>2/6 (33%)</td>
<td>6/6 (100%)</td>
<td>5/9 (56%)</td>
<td>9/9 (100%)</td>
</tr>
<tr>
<td>11</td>
<td>1/1 (100%)</td>
<td>1/1 (100%)</td>
<td>3/8 (38%)</td>
<td>7/8 (87%)</td>
<td>4/9 (44%)</td>
<td>8/9 (89%)</td>
</tr>
<tr>
<td>12</td>
<td>0/0</td>
<td>0/0</td>
<td>3/9 (33%)</td>
<td>9/9 (100%)</td>
<td>3/9 (33%)</td>
<td>9/9 (100%)</td>
</tr>
<tr>
<td>13</td>
<td>0/0</td>
<td>0/0</td>
<td>8/17 (47%)</td>
<td>17/17 (100%)</td>
<td>8/17 (47%)</td>
<td>17/17 (100%)</td>
</tr>
<tr>
<td>14</td>
<td>0/0</td>
<td>0/0</td>
<td>3/5 (60%)</td>
<td>5/5 (100%)</td>
<td>3/5 (60%)</td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>15</td>
<td>0/0</td>
<td>0/0</td>
<td>2/5 (40%)</td>
<td>5/5 (100%)</td>
<td>2/5 (40%)</td>
<td>5/5 (100%)</td>
</tr>
<tr>
<td>16</td>
<td>0/0</td>
<td>0/0</td>
<td>0/2 (0%)</td>
<td>2/2 (100%)</td>
<td>0/2 (0%)</td>
<td>2/2 (100%)</td>
</tr>
<tr>
<td>17</td>
<td>0/0</td>
<td>0/0</td>
<td>0/1 (0%)</td>
<td>1/1 (100%)</td>
<td>0/1 (0%)</td>
<td>1/1 (100%)</td>
</tr>
<tr>
<td>18</td>
<td>0/0</td>
<td>0/0</td>
<td>2/2 (100%)</td>
<td>2/2 (100%)</td>
<td>2/2 (100%)</td>
<td>2/2 (100%)</td>
</tr>
</tbody>
</table>

*Ceiling of care was not clearly documented for 4 patients

Failure Assessment scores with pre-specified exclusion criteria and reverse triage at 48 hours. During H1N1 influenza seasons, the performance of the OHPIP was evaluated in several studies. In one study of two UK centres, during the 2010-2011 influenza epidemic, the OHPIP predicted critical care admission with a sensitivity of 76% and a specificity of 91%. However, the reverse triage component at 48 hours poorly predicted non-survival. In that study, seven out of nine patients who would have been identified for palliation by the OHPIP ended up surviving to hospital discharge.
Figure 3. Receiver-operating curves for age, Clinical Frailty Scale score (CFS), or Decision Support Tool (DST) for prediction of either mortality or a ceiling of treatment not including intubation and ventilation. Comparing area under the curve (AUC) values, the decision support tool was more discriminative than age alone ($p<0.001$), and was potentially more discriminative than CFS ($p=0.08$) when predicting not for intubation and ventilation (non-I&V) treatment ceiling. All other pairwise comparisons yielded larger $p$-values (Appendix Figure 4).

Figure 4. Area Under Curve (AUC) estimates for each combination of predictor and outcome. Error bars show 95% bootstrap confidence intervals.
In our study, both the Clinical Frailty Score (with a cut-off of ≥5) and the DST had relatively low specificity for predicting mortality (69% and 63%). However, that may reflect the mixed illness severity. A number of patients with higher scores of CFS or DST never developed severe respiratory failure. However, both appeared to be highly associated with treatment limitation decisions. CFS and the DST had 99% and 96% specificity for treatment limitation. Among patients who were reviewed by intensive care, there were no patients with a CFS of ≥5 who survived. There were two patients with a DST above the threshold suggested in the proposed tool who were admitted to intensive care; one of these patients survived to discharge. There were two patients with lower DST, whose ceiling of care was capped at NIV. Both cases were medically complex in ways not captured by any of the elements of the DST.

Clinical frailty has been proposed as a tool for intensive care triage because of its relationship with outcome for critically ill older patients. A 2017 meta-analysis found that frailty was associated with higher in-hospital and long-term mortality. In the “VIP1” (Very old Intensive Care Patients) study, 5021 patients aged 80 years or over were followed up after admission to intensive care units in 21 European countries. There was a linear relationship between the clinical frailty score and mortality. The authors observed about a 30% 30-day mortality for patients with a CFS of 3 requiring emergency intensive care admission, rising to 75% mortality for CFS 9. However, the absolute rate of mortality in older frail patients admitted to intensive care varies between studies.

While the use of clinical frailty in COVID-19 triage has been defended, there are other factors that are relevant to decisions about the benefit and appropriateness of intensive care. Co-morbidity has been clearly linked to mortality in patients with COVID-19, and is a risk factor for length of stay in intensive care. Age also appears to be an independent risk factor for intensive care outcome. Younger patients who are clinically frail also have increased intensive care mortality. In a Canadian cohort of about 200 patients aged 50–65 years, frailty was linked with increased mortality risk over ensuing 12 months in adjusted analysis. The authors concluded that the probability of survival is higher for younger compared with older frail patients, but their risk is age-shifted compared to peers – i.e. they have a mortality risk equivalent to an older chronological age.

There are some limitations to our study. The surge in Oxford and the United Kingdom as a whole was significantly less than anticipated and the hospitals’ increased capacity through reorganisation of services was sufficient to offer intubation and ventilation to patients using the same or very similar standards as under non-COVID-19 circumstances. Intensive care and medical doctors were thus not faced with decisions to ration care in the way that was predicted. The numbers of patients in this single centre study are relatively small, and triage scores were evaluated retrospectively from medical records, which may have introduced inaccuracy or bias in classification. A small proportion did not have clear documentation of ceiling of treatment. The decision-making and outcomes reported in this study may be different from those in other settings.

**Figure 5.** Estimated sensitivities and specificities for each predictor and outcome at default cut-offs.
in other centres with greater pressure on intensive care, or indeed from those in our own centre at other time points during the pandemic. The mortality rate was lower than average for the UK (overall, in the first wave there were 10,624 UK ICU admissions with COVID-19, of which 6232 were discharged, 4109 died – a mortality rate of 38.7%)\(^{19}\).

Since the clinical frailty score was included in NICE guidance relating to COVID-19 management, it is possible that knowledge of CFS directly influenced decisions about treatment and therefore the calculated sensitivity/specificity may not be accurate. (However, in half of the cohort, the CFS was not documented in the medical record, potentially meaning that at least in that group it did not directly influence decisions). Since treatment limitation decisions may lead to higher mortality, the relationship between the DST and mortality is potentially influenced by self-fulfilling prophecies\(^{10}\). However, in our cohort the DST did not appear to be more predictive of mortality than either frailty or age alone. Calculated values for sensitivity and specificity obtained from a case-control study would have generated more accurate assessment.

Hewitt et al. found greater correlation between CFS score and mortality in their multi-centre study\(^{11}\). Since this decision-support tool was proposed, other prognostic risk stratification scores for patients with COVID-19 have been assessed in large cohorts e.g. ISARIC (International Severe Acute Respiratory and Emerging Infection Consortium)\(^{12}\).

**Conclusions**

To our knowledge, this is the first study to formally evaluate the potential performance of the proposed MEAG decision support tool. While the DST was not used, it was a strong predictor of clinical decisions made by the doctors in the OUH hospitals; in our patient cohort it appeared to have better discriminatory power than age alone, and potentially better power than the clinical frailty score. When it comes to predicting mortality, however, the three metrics we considered performed similarly. This data may be of use in the assessment of other triage tools in this or future pandemics.

**Data availability**

Underlying data

Open Science Framework: COVID triage data Decision Support Tool. https://doi.org/10.17605/OSF.IO/K2FTW\(^{15}\).

Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

**Acknowledgments**

We thank CE Richards and C Turberfield for their advice.

---

**References**

   [PubMed Abstract] [Publisher Full Text]

   [PubMed Abstract] [Publisher Full Text] [Free Full Text]

   [Publisher Full Text]

   [PubMed Abstract] [Publisher Full Text]

   [PubMed Abstract] [Publisher Full Text]

   [PubMed Abstract] [Publisher Full Text]

   [Reference Source]

   [PubMed Abstract] [Publisher Full Text] [Free Full Text]

   [Reference Source]

    [PubMed Abstract] [Publisher Full Text]

    [Reference Source]

    [PubMed Abstract] [Publisher Full Text] [Free Full Text]

    [Publisher Full Text]

    [PubMed Abstract] [Publisher Full Text]


    [PubMed Abstract] [Publisher Full Text] [Free Full Text]

    [PubMed Abstract] [Publisher Full Text] [Free Full Text]

    [PubMed Abstract] [Publisher Full Text]

    [PubMed Abstract] [Publisher Full Text]

    [PubMed Abstract] [Publisher Full Text] [Free Full Text]

    [PubMed Abstract] [Publisher Full Text]

22. Darvall JN, Bellomo R, Paul E, et al.: Frailty in very old critically ill patients in...


