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: 2 approved with reservations]

Emily Simon Thomas ¹, Bryony Peiris ¹, Leon Di Stefano²,
Matthew J. Rowland^{1,3}, Dominic Wilkinson ^{1,4,5}

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

V1 First published: 19 Jul 2021, 6:183
<https://doi.org/10.12688/wellcomeopenres.16939.1>

Latest published: 19 Jul 2021, 6:183
<https://doi.org/10.12688/wellcomeopenres.16939.1>

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.

Open Peer Review

Approval Status ? ?

	1	2
version 1 19 Jul 2021	view	view
1. Edy Y Kim , Brigham and Women's Hospital, Harvard Medical School, Boston, USA		
2. Magnolia Cardona , Gold Coast University Hospital, Southport, Australia Bond University, Gold Coast, Australia		
Any reports and responses or comments on the article can be found at the end of the article.		

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. *The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

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: 2 approved with reservations]** 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^{1,2}. Pandemic modelling suggested that rates of infection in the UK would exceed ICU capacity³.

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^{4,5}. 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 way⁶.

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, 2020⁷. It recommended that all admissions be assessed for frailty, using

the Clinical Frailty Scale (CFS)⁸. 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 benefit⁷. 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 age, frailty and co-morbidities (Figure 1)^{9,10}. 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 officials¹¹, 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 care¹⁰. Secondly, such scores

COVID-19 DECISION SUPPORT TOOL

1

AGE	POINTS
<50	0
50-60	1
61-65	2
66-70	3
71-75	4
76-80	5
>80	6

2 Clinical Frailty Scale*

1 Very Fit – People who are robust, active, energetic and motivated. These people commonly exercise regularly. They are among the fittest for their age.

2 Well – People who have no active disease symptoms but are less fit than category 1. Often, they exercise or are very active occasionally, e.g. seasonally.

3 Managing Well – People whose medical problems are well controlled, but are not regularly active beyond routine walking.

4 Vulnerable – While not dependent on others for daily help, often symptoms limit activities. A common complaint is being “slowed up”, and/or being tired during the day.

5 Mildly Frail – These people often have more evident slowing, and need help in high order IADLs (finances, transportation, heavy housework, medications). Typically, mild frailty progressively impairs shopping and walking outside alone, meal preparation and housework.

6 Moderately Frail – People need help with all outside activities and with keeping house. Inside, they often have problems with stairs and need help with bathing and might need minimal assistance (cuing, standby) with dressing.

7 Severely Frail – Completely dependent for personal care, from whatever cause (physical or cognitive). Even so, they seem stable and not at high risk of dying (within ~ 6 months).

8 Very Severely Frail – Completely dependent, approaching the end of life. Typically, they could not recover even from a minor illness.

9 Terminally Ill – Approaching the end of life. This category applies to people with a life expectancy <6 months, who are not otherwise evidently frail.

* 1. Canadian Study on Health & Aging, Revised 2008.
2. K. Rockwood et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005; 173: 489-495.

3

CO-MORBIDITY	POINTS
In last 3 years, cardiac arrest from any cause	2
Chronic condition causing:	
• ≥3 hospital admissions in the last year	2
• ≥4 weeks continuous admission for current inpatients	2
Congestive heart failure with symptoms at rest or on minimal exertion	1
Chronic lung disease with symptoms at rest or on minimal exertion	1
Hypertension	1
Severe and irreversible neurological condition including dementia	1
Chronic Liver Disease with Child-Pugh score ≥ 7	1
End stage chronic renal failure requiring renal replacement therapy	1
Diabetes mellitus requiring medication	1
Uncontrolled or active malignancy	1

TOTAL = SUM OF THE 3 DOMAINS ABOVE (-1 FOR FEMALE SEX)

! There may be situations arising that are outside the scope of the framework that require special consideration, thus clinical discretion will continue to apply. Frailty scoring is used as a proxy for physiological frailty which leads to reduced chances of recovery in ICU, therefore where conditions pre-exist impact on physical activity but are stable and inappropriately affect the score, then that situation requires special consideration.

POINTS	TREATMENT	FAILURE OF FIRST LINE MANAGEMENT	NOTES
Group 1 ≤ 8	ICU-based care	Palliation or ECMO	Usual criteria for ECMO and <60 years
Group 2 > 8	Ward-based care	Step 3	Consider trial of CPAP
Group 3 Patients not normally for full active management or failed CPAP trial	Facemask oxygen	Palliation	Consider domiciliary care

Deviations from ARDS guideline	Investigations	Support	Treatment
Step 1 ≤ 8	Tracheo-bronchial aspirate for respiratory viruses. Avoid CT & bronchoscopy unless indicated. H score screen blood tests, D-dimers, LDH & troponin (all days) Lung US to reduce X-ray usage	CPAP trial in ICU or with rapid access to intubation (for hours not days) Avoid HFNO	CAP antimicrobials Continue single agent prophylaxis in +ve pts Disease modifying agents as part of RCT
Step 2 > 8	Standard swabs	Ward-based CPAP	CAP antimicrobials Continue single agent prophylaxis in +ve pts
Step 3 Patients not normally for full active management or failed CPAP trial	Standard swabs	Facemask oxygen	CAP antimicrobials Continue single agent prophylaxis in +ve pts

Figure 1. Draft coronavirus disease 2019 (COVID-19) decision-support tool⁹.

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)⁹. 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 ≤ 8 vs > 8 ; 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 pROC¹². We computed sensitivity and specificity values for ROC curves using that package's `roc` function. To generate confidence intervals (CIs) for sensitivity and specificity of each predictor, we used the package's bootstrap intervals¹³. 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 test¹⁴.

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)¹⁵.

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 <35 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 cannulae 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.
IQR=interquartile range.

	ICU reviewed (n=75)	Non-reviewed (n=76)	Total (n=151)	Test of no difference between groups
Age: mean (±SD, range)	59 (±14.02, 23-83)	76 (±13.8, 39-96)	67 (±16.2, 23-96)	$p = 5.2 \times 10^{-10}$ (Mann-Whitney test using age categories)
Gender	31F:44M	34F:42M	65F:86M	$p = 0.80$ ($\chi^2 = 0.07$)
Clinical Frailty Score: median IQR (Q1-Q3)	2 1 (2-3)	5 2 (4-6)	3 2.5 (3-5.5)	$p = 1.4 \times 10^{-15}$ (Mann-Whitney)
Decision Support Tool Score: median IQR (Q1-Q3)	5 4 (3-7)	12 4 (9-13)	8 8 (4-12)	$p = 2.6 \times 10^{-16}$ (Mann-Whitney)

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-.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 pandemic¹⁶. It combines Sequential Organ

Table 2. Logistical regressions of mortality on clinical frailty (CFS), Decision Support Tool (DST) score, age.

Predictor of mortality	Likelihood ratio test from logistic regression
DST score	$p = 1.4 \times 10^{-5}$ ($\chi^2 = 18.8$ on 1 degree of freedom)
Clinical frailty score	$p = 2.8 \times 10^{-4}$ ($\chi^2 = 13.2$ on 1 degree of freedom)
Age (as a discrete factor)	$p = 6.6 \times 10^{-4}$ ($\chi^2 = 23.4$ on 6 degrees of freedom)

Table 3. Mortality and treatment limitation per clinical frailty (CFS) and Decision Support Tool (DST) scores.
ICU=intensive care unit.

	ICU (n=75)		Non-ICU (n=76)		Total (n=151)	
CFS	mortality	treatment limitation	mortality	treatment limitation	mortality	treatment limitation
1	0/5 (0%)	0/5 (0%)	0/3 (0%)	0/3 (0%)	0/9 (0%)	0/9 (0%)
2	6/35 (17%)	1/34* (3%)	2/3 (50%)	1/3 (25)	8/36 (22%)	4/35* (11%)
3	6/26 (23%)	8/25* (32%)	1/11 (9%)	3/9* (33%)	8/38 (21%)	11/35* (31%)
4	1/5 (20%)	2/5 (40%)	0/4 (0%)	2/4 (50%)	1/9 (11%)	4/9 (44%)
5	1/1 (100%)	0/1 (0%)	5/20 (25%)	20/20 (100%)	6/21 (29%)	20/21 (95%)
6	3/3 (100%)	3/3 (100%)	7/17 (41%)	17/17 (100%)	11/20 (55%)	20/20 (100%)
7	0/0	0/0	8/16 (50%)	16/16 (100%)	8/16 (50%)	16/16 (100%)
8	0/0	0/0	0/1 (0%)	1/1 (100%)	0/1 (0%)	1/1 (100%)
9	0/0	0/0	1/1 (100)	1/1 (100%)	1/1 (100%)	1/1 (100%)
DST						
0	0/1 (0%)	0/1 (0%)	0/1 (0%)	0/1 (0%)	0/2 (0%)	0/2 (0%)
1	0/5 (0%)	0/5 (0%)	0/0	0/0	0/5 (0%)	0/5 (0%)
2	0/9 (0%)	0/8* (0%)	0/1 (0%)	0/1 (0%)	0/11 (0%)	0/10* (0%)
3	1/12 (8%)	0/12 (0%)	0/3 (0%)	0/2* (0%)	1/13 (8%)	0/12* (0%)
4	1/10 (10%)	1/10 (10%)	0/1 (0%)	0/1 (0%)	1/12 (7%)	1/12 (7%)
5	2/10 (20%)	2/10 (18%)	0/4 (0%)	0/4 (0%)	2/14 (12%)	2/14 (12%)
6	1/5 (20%)	1/5 (20%)	2/3 (67%)	1/3 (33%)	3/8 (38%)	2/8 (25%)
7	4/7 (57%)	2/7 (29%)	0/1 (0%)	0/0*	4/8 (50%)	2/7* (29%)
8	4/7 (57%)	3/6* (50%)	0/3 (0%)	2/3 (67%)	4/10 (40%)	5/9* (56%)
9	1/5 (20%)	3/5 (60%)	0/4 (0%)	4/4 (100%)	1/9 (11%)	7/9 (78%)
10	3/3 (100%)	3/3 (100%)	2/6 (33%)	6/6 (100%)	5/9 (56%)	9/9 (100%)
11	1/1 (100%)	1/1 (100%)	3/8 (38%)	7/8 (87%)	4/9 (44%)	8/9 (89%)
12	0/0	0/0	3/9 (33%)	9/9 (100%)	3/9 (33%)	9/9 (100%)
13	0/0	0/0	8/17 (47%)	17/17 (100%)	8/17 (47%)	17/17 (100%)
14	0/0	0/0	3/5 (60%)	5/5 (100%)	3/5 (60%)	5/5 (100%)
15	0/0	0/0	2/5 (40%)	5/5 (100%)	2/5 (40%)	5/5 (100%)
16	0/0	0/0	0/2 (0%)	2/2 (100%)	0/2 (0%)	2/2 (100%)
17	0/0	0/0	0/1 (0%)	1/1 (100%)	0/1 (0%)	1/1 (100%)
18	0/0	0/0	2/2 (100%)	2/2 (100%)	2/2 (100%)	2/2 (100%)

*Ceiling of care was not clearly documented for 4 patients

Failure Assessment scores with pre-specified exclusion criteria and reverse triage at 48 hours¹⁶. During H1N1 influenza seasons, the performance of the OHPIP was evaluated in several studies^{17,18}. 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%¹⁹. 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-l&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.

Figure 5. Estimated sensitivities and specificities for each predictor and outcome at default cut-offs.

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 mortality²⁰. 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 countries²¹. 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^{22–25}.

While the use of clinical frailty in COVID-19 triage has been defended¹⁰, 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^{26,27}. Age also appears to be an independent risk factor for intensive care outcome^{26,27}. 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 analysis²⁸. 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 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%)²⁹.

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 prophecies³⁰. 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 cohort may be different from an unselected cohort. A larger prospective multi-centre study from a multi-organisational study would have generated more accurate assessment. Hewitt *et al.* found greater correlation between CFS score and mortality in their multi-centre study³¹. 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)³².

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

- Wilkinson D: **ICU triage in an impending crisis: uncertainty, pre-emption and preparation.** *J Med Ethics.* 2020; **46**(5): 287–8. [PubMed Abstract](#) | [Publisher Full Text](#)
- Verelst F, Kuylen E, Beutels P: **Indications for healthcare surge capacity in European countries facing an exponential increase in coronavirus disease (COVID-19) cases, March 2020.** *Euro Surveill.* 2020; **25**(13): 2000323. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Ferguson N, Laydon D, Nedjati-Gilani G, *et al.*: **Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand.** Imperial College London, 2020; (accessed 8 Sep 2020). [Publisher Full Text](#)
- Strauss MJ, LoGerfo JP, Yeltatzie JA, *et al.*: **Rationing of intensive care unit services. An everyday occurrence.** *JAMA.* 1986; **255**(9): 1143–6. [PubMed Abstract](#) | [Publisher Full Text](#)
- Joynt GM, Gomersall CD, Tan P, *et al.*: **Prospective evaluation of patients refused admission to an intensive care unit: triage, futility and outcome.** *Intensive Care Med.* 2001; **27**(9): 1459–65. [PubMed Abstract](#) | [Publisher Full Text](#)
- Emanuel EJ, Persad G, Upshur R, *et al.*: **Fair Allocation of Scarce Medical Resources in the Time of Covid-19.** *N Engl J Med.* 2020; **382**(21): 2049–2055. [PubMed Abstract](#) | [Publisher Full Text](#)
- NICE: **COVID-19 rapid guideline: critical care in adults.** NICE, 2020; (accessed 8 Sep 2020). [Reference Source](#)
- Rockwood K, Song X, MacKnight C, *et al.*: **A global clinical measure of fitness and frailty in elderly people.** *CMAJ.* 2005; **173**(5): 489–95. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Foster P, Staton B, Rovnick N: **NHS 'score' tool to decide which patients receive critical care.** *Financial Times*, 2020; (accessed 8 Sep 2020). [Reference Source](#)
- Wilkinson DJC: **Frailty Triage: Is Rationing Intensive Medical Treatment on the Grounds of Frailty Ethical?** *Am J Bioeth.* 2020; 1–22. [PubMed Abstract](#) | [Publisher Full Text](#)
- Kirkpatrick DD, Mueller B: **U.K. Backs Off Medical Rationing Plan as Coronavirus Rages.** *N.Y. Times.*, 2020; (accessed 8 Sep 2020). [Reference Source](#)
- Robin X, Turck N, Hainard A, *et al.*: **pROC: an open-source package for R and S+ to analyze and compare ROC curves.** *BMC Bioinformatics.* 2011; **12**(1): 77. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Fawcett T: **An introduction to ROC analysis.** *Pattern Recognit Lett.* 2006; **27**(8): 861–874. [Publisher Full Text](#)
- DeLong ER, DeLong DM, Clarke-Pearson DL: **Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach.** *Biometrics.* 1988; **44**(3): 837–845. [PubMed Abstract](#) | [Publisher Full Text](#)
- Thomas ES: **COVID triage data Decision Support Tool.** 2021. <http://www.doi.org/10.17605/OSF.IO/K2FTW>
- Christian MD, Hawryluck L, Wax RS, *et al.*: **Development of a triage protocol for critical care during an influenza pandemic.** *CMAJ.* 2006; **175**(11): 1377–81. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Adeniji KA, Cusack R: **The Simple Triage Scoring System (STSS) successfully predicts mortality and critical care resource utilization in H1N1 pandemic flu: a retrospective analysis.** *Crit Care.* 2011; **15**(1): R39. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Cheung WK, Myburgh J, Seppelt IM, *et al.*: **A multicentre evaluation of two intensive care unit triage protocols for use in an influenza pandemic.** *Med J Aust.* 2012; **197**(3): 178–181. [PubMed Abstract](#) | [Publisher Full Text](#)
- Morton B, Tang L, Gale R: **Performance of influenza-specific triage tools in an H1N1-positive cohort: P/F ratio better predicts the need for mechanical ventilation and critical care admission.** *Br J Anaesth.* 2015; **114**(6): 927–933. [PubMed Abstract](#) | [Publisher Full Text](#)
- Muscudere J, Waters B, Varambally A, *et al.*: **The impact of frailty on intensive care unit outcomes: a systematic review and meta-analysis.** *Intensive Care Med.* 2017; **43**(8): 1105–1122. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
- Flaatten H, De Lange DW, Morandi A, *et al.*: **The impact of frailty on ICU and 30-day mortality and the level of care in very elderly patients (≥ 80 years).** *Intensive Care Med.* 2017; **43**(12): 1820–1828. [PubMed Abstract](#) | [Publisher Full Text](#)
- Darvall JN, Bellomo R, Paul E, *et al.*: **Frailty in very old critically ill patients in**

- Australia and New Zealand: a population-based cohort study.** *Med J Aust.* 2019; **211**(7): 318–23.
[PubMed Abstract](#) | [Publisher Full Text](#)
23. Muessig JM, Nia AM, Masyuk M, *et al.*: **Clinical Frailty Scale (CFS) reliably stratifies octogenarians in German ICUs: a multicentre prospective cohort study.** *BMC Geriatr.* 2018; **18**(1): 162.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 24. Simpson A, Puxty K, McLoone P, *et al.*: **Comorbidity and survival after admission to the intensive care unit: A population-based study of 41,230 patients.** *J Intensive Care Soc.* 2021; **22**(2): 143–151.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 25. Nielsen AB, Thorsen-Meyer HC, Belling K, *et al.*: **Survival prediction in intensive-care units based on aggregation of long-term disease history and acute physiology: a retrospective study of the Danish National Patient Registry and electronic patient records.** *Lancet Digit Health.* 2019; **1**(2): e78–89.
[PubMed Abstract](#) | [Publisher Full Text](#)
 26. Cardoso FS, Papoila AL, Machado RS, *et al.*: **Age sex, and comorbidities predict ICU admission or mortality in cases with SARS-CoV2 infection: a population-based cohort study.** *Crit Care.* 2020; **24**(1): 465.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 27. Yang Y, Yang KS, Hsann YM, *et al.*: **The effect of comorbidity and age on hospital mortality and length of stay in patients with sepsis.** *J Crit Care.* 2010; **25**(3): 398–405.
[PubMed Abstract](#) | [Publisher Full Text](#)
 28. Bagshaw M, Majumdar SR, Rolfson DB, *et al.*: **A prospective multicenter cohort study of frailty in younger critically ill patients.** *Crit Care.* 2016; **20**(1): 175.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 29. ICNARC: **ICNARC report on COVID-19 in critical care 31 July 2020.** ICNARC Case Mix Programme Database.
[Reference Source](#)
 30. Wilkinson D: **The self-fulfilling prophecy in intensive care.** *Theor Med Bioeth.* 2009; **30**(6): 401–10.
[PubMed Abstract](#) | [Publisher Full Text](#)
 31. Hewitt J, Carter B, Vilchez-Moraga A, *et al.*: **The effect of frailty on survival in patients with COVID-19 (COPE): a multicentre, European, observational cohort study.** *Lancet Public Health.* 2020; **5**(8): E444–E451.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
 32. Knight SR, Ho A, Pius R, *et al.*: **Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score.** *BMJ.* 2020; **370**: m3339.
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)

Open Peer Review

Current Peer Review Status: ? ?

Version 1

Reviewer Report 07 February 2022

<https://doi.org/10.21956/wellcomeopenres.18692.r48114>

© 2022 Cardona M. This is an open access peer review report 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.

? **Magnolia Cardona**

¹ Evidence-Based Practice Professorial Unit, Gold Coast University Hospital, Southport, Qld, Australia

² Institute for Evidence Based Healthcare, Bond University, Gold Coast, Australia

The topic of whether a proposed triage tool would have performed adequately in selecting the right patients for access to ICU resources is of high relevance and importance as the COVID-19 crisis continues.

However, it is unfortunate that the tool was not formally implemented and the study was conducted retrospectively, as it precluded accurate and complete collection of frailty CFS score, and integral component of the DST. The fact that assumptions had to be made based on inconsistent case notes -even if consensus achieved by assessors- leads to a major flaw in the interpretation of results. The extreme CFS scores of 1 and >7 are not the problem. It is the ascertainment of the 4-5 scores which are around the proposed threshold that counts.

Importantly, there would be no practical reason to analyse the predictive ability of age alone, or frailty alone (even in light of the NICE suggestion of less uncertainty about benefit for the frail), as no triage tool can ethically justify exclusion of patients from intensive management on the basis of a single parameter, even in crisis times. (Journal of Critical care, December 2021 <https://www.sciencedirect.com/science/article/pii/S0883944121001726>¹)

The shape of the single item ROC curves confirms that it is hardly worth attempting this prediction.

Next, what was the justification for the timeframe of data collection? It was not a preliminary assessment of the accuracy of prediction to enable urgent implementation.

The world witnessed plenty of cases in the UK and yet the sample was too small for meaningful predictive analysis and taken from such a narrow timeframe one month for the ICU patients and 2 weeks for the non-ICU control sample, that it raises concerns about the representativeness, comparability between samples and soundness of the methods. While the use of ROC curves and sensitivity/specificity estimations are generally appropriate, the language used in this

retrospective analysis should have been about the % classified correctly, or the hypothetical tool performance rather than 'predicted ceilings of care'.

Overall the study had a worthwhile objective but the retrospective nature of the analysis will have its inherent drawbacks; I believe a larger sample over a longer period including only people with complete information (avoiding assumptions) to confirm or disprove the tentative conclusions is warranted and feasible in the UK health system.

References

1. Cardona M, Dobler CC, Koreshe E, Heyland DK, et al.: A catalogue of tools and variables from crisis and routine care to support decision-making about allocation of intensive care beds and ventilator treatment during pandemics: Scoping review. *J Crit Care*. **66**: 33-43 [PubMed Abstract](#) | [Publisher Full Text](#)

Is the work clearly and accurately presented and does it cite the current literature?

Partly

Is the study design appropriate and is the work technically sound?

Partly

Are sufficient details of methods and analysis provided to allow replication by others?

Partly

If applicable, is the statistical analysis and its interpretation appropriate?

Partly

Are all the source data underlying the results available to ensure full reproducibility?

No source data required

Are the conclusions drawn adequately supported by the results?

Partly

Competing Interests: Only potentially perceived competing interests, which I consider confluence of interest instead. I have researched this area and published before. My reference is in the report. I derive no financial or other benefit whatsoever from that publication

Reviewer Expertise: Gerontology and geriatrics, decision-making in healthcare, health systems research

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 06 January 2022

<https://doi.org/10.21956/wellcomeopenres.18692.r47649>

© 2022 Kim E. This is an open access peer review report 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.

Edy Y Kim

Division of Pulmonary and Critical Care Medicine, Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

The authors perform a single-center, retrospective cohort study (N=151 in ICU, non-ICU subgroups) to evaluate the discrimination of age, Frailty scale and decision-support tool (DST, from UK's MEAG) for mortality and treatment ceiling decisions (categorized as MV; NIPPV or HFNC; or less than NIPPV/HFNC). Age, frailty, DST associated with mortality/tx limitation. DST >8 with sens/spec 65/63 and 80/96 for mortality and tx limitation, respectively and outperformed age alone but not frailty (although a trend c/w frailty).

Strengths:

1. The authors address an important question, the performance of "clinical triage" algorithms, that is relatively understudied.
2. The authors examine clinicians' decisions on treatment limitation in the absence of crisis standards of care. This is an important endpoint that is understudied compared to mortality or WHO-type ordinal scales of respiratory failure.
3. Clear, concise writing that communicated UK-/center-specific details of clinical care.

Major weaknesses:

1. Per authors, frailty score was documented in ~half of patients. The authors do note this as a study limitation. While possibly limited by their N, a sensitivity analysis of the effect of frailty scores documented during hospital admission might be helpful.
2. At what point during hospital admission were treatment limitation decisions recorded (e.g., at admission? During decompensation? At the point of triage to ICU or not)? The time point of treatment limitation decisions could introduce bias (e.g., if many tx limitation decisions were made during clinical deterioration, a highly frail patient that would have been scored as "+treatment limitation" during deterioration might not have a treatment limitation recorded, if this frail patient remained "well" throughout the admission).

Minor comments:

1. Define ITU at first use (not standard term in some countries like US).
2. Inherent limitations of a single cohort/single-center retrospective study of moderate N (noted by authors as a limitation).
3. It is outside of the scope of this manuscript, but an interesting future direction could be a comparison of crisis standards of care (CSC) algorithms that incorporate parameters from clinical presentation (e.g., SOFA as in some U.S. state CSC guidelines) on prediction of treatment limitation decisions.

Is the work clearly and accurately presented and does it cite the current literature?

Yes

Is the study design appropriate and is the work technically sound?

Yes

Are sufficient details of methods and analysis provided to allow replication by others?

Yes

If applicable, is the statistical analysis and its interpretation appropriate?

Yes

Are all the source data underlying the results available to ensure full reproducibility?

No source data required

Are the conclusions drawn adequately supported by the results?

Yes

Competing Interests: E.Y.K. is a member of the advisory board for Cell Reports Medicine. E.Y.K. is a Managing Editors for COVIDprotocols.org. E.Y.K. is a co-investigator and receives no financial remuneration from NCT04389671 (Windtree Therapeutics) testing lucinactant (surfactant-like treatment) in COVID-19 patients. E.Y.K. is a member of the Steering Committees for and receives no financial remuneration from NCT04409834 (Prevention of arteriovenous thrombotic events in critically ill COVID-19 patients, TIMI group) and REMAP-CAP ACE2 renin-angiotensin system (RAS) modulation domain. E.Y.K. has received unrelated research funding from Bayer AG, Roche Pharma Research and Early Development, the U.S. National Institutes of Health, the American Heart Association, the American Lung Association, the American Thoracic Society, and the Brigham and Women's Hospital Department of Medicine.

Reviewer Expertise: Clinical / translational in critical care, respiratory failure, COVID-19. Translational / basic in immunology.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.