TRIAGING COVID-19 IN THE EMERGENCY DEPARTMENT

An Evidence Review from Penn Medicine's Center for Evidence-based Practice

February 2021

Project director:.....Nikhil K. Mull, MD (CEP)
Lead analyst: .....Matthew D. Mitchell, PhD (CEP)
Clinical review: .....Eric Bressman, MD (Medicine); S. Ryan Greysen, MD, MHS, MA, FHM (CEP); Kathleen Lee, MD (PPMC); Anna U. Morgan, MD (Medicine)

Keywords: COVID-19, Emergency Department, triage

EVIDENCE SUMMARY

- Several published algorithms for appraising disease severity in patients with COVID-19 disease or determining whether they can be safely discharged from the Emergency Department (or both) were identified. None of these algorithms have been validated in clinical studies measuring their impact on patient outcomes.
- Among those algorithms, there is consensus that patients with low oxygen saturation should not be discharged (thresholds range from oxygen saturation of 91% to 94% without supplemental oxygen). Some of the algorithms also include a test for eliciting oxygen desaturation on exertion (walking test). Most algorithms also include respiratory rate as a criterion for patient discharge. Other vital signs are considered in some but not all algorithms, and algorithms differ in which comorbidities and other risk factors should be considered in triage decisions. A table of representative algorithms is presented on page 14.
- Many different biomarkers and other laboratory test results have been shown to be significantly associated with adverse out–comes in COVID-19 disease. There is not sufficient evidence to determine which measure has the greatest predictive value.
- Several composite risk scoring systems including vital signs and/or test results have been proposed for use in management of COVID-19 patients. None of them have been validated in clinical studies measuring their impact on patient outcomes. The 4C score, the quick COVID Severity Index (qCSI), and the NEWS2 score (which is not COVID-specific) have been spoken of favorably in review articles. Evidence from studies directly comparing their effectiveness is lacking. Studies of their predictive ability are made weaker by use of area under the ROC as an outcome measure and having been performed early in the pandemic, when adverse outcomes were more common. We conclude that there is moderate-strength evidence that these tools have some predictive ability, but we cannot draw conclusions about their comparative effectiveness or clinical utility.

CENTER for EVIDENCE-BASED PRACTICE

Table of Contents

Penn Medicine

Introduction

The pandemic caused by the SARS-CoV-2 coronavirus (COVID-19) has caused a great increase in patients presenting to hospital emergency rooms and triage centers with respiratory symptoms consistent with the coronavirus infection. Some of these patients have mild disease that can be managed safely at home, but others may suffer a significant deterioration of their condition and require inpatient or critical care unit treatment. Effective triage of these patients is essential to maintaining medical center operations during the pandemic.

The purpose of this report is to identify pathways, algorithms, and clinical risk prediction tools that can be used in Emergency Department triage of patients with suspected or confirmed COVID-19 disease and to review the evidence on their effectiveness. There are numerous pathways, algorithms, and scales for use in patients who already have been admitted as hospital inpatients; management of these patients is outside the scope of this report.

Previous CEP Reports

Please consult the CEP web site for a complete catalog of reports relating to COVID-19 disease. All COVID-19-related reports are freely downloadable from the site.

Methods

PROTOCOL FOR SYSTEMATIC REVIEW

SPECIFIC AIM:

Identify and summarize clinical pathways, algorithms, and risk scoring scales for patients with suspected or confirmed COVID-19 disease presenting to the Emergency Department or in a short-stay observation unit.

METHODS:

Inclusion and exclusion criteria:

Participants: Adult patients discharged from the hospital.

Interventions: Use of a clinical pathway, algorithm, risk scoring tool, or other system for ascertaining whether patients can safely be discharged to home, whether they should remain under short-term observation (less than 48 hours), or whether they should be admitted to the hospital. Tools intended to predict the likelihood that the patient has COVID-19 disease are outside the scope of this report. Tools intended for use in patients who have already been admitted to the hospital are outside the scope of this report.

Comparisons: All comparisons, including usual care.

Outcomes: All outcomes. Key outcomes include need for readmission to the hospital, length of stay, intubation, transfer to ICU, mortality.

Timing: Risk assessment or pathway applied on initial presentation (within 4 hours or repeated assessment 24-48 hours after initial). Outcomes from initial presentation to 30 days.

Setting: Emergency department or dedicated COVID-19 intake/triage center, hospital-based observation unit. Outpatient clinics are outside the scope of this report.

Other: Limit to studies carried out in OECD countries.

Data collection

Databases: NICE Evidence Search, ECRI Guidelines Trust, Medline, EMBASE.
NOTE: additional COVID-specific guidance sources will be searched as per the methods for a CEP Rapid Guidance Summary.

Study design: All comparative studies, including pre-post designs

Study quality assessment: The draft CEP appraisal tool for clinical pathways is presented in Appendix A. Other CEP standard review methods, including scales for quality assessment of guidelines, systematic reviews, and primary studies can be found in the Methods section of the CEP web site. (www.uphs.upenn.edu/cep/methods).

Data synthesis (calculation of relative risks and confidence intervals, meta-analyses, exploration of heterogeneity):
Random-effects meta-analysis following Cochrane methods if quantity and homogeneity of data permit, otherwise qualitative analysis.

Assessment of quality of evidence base: GRADE.

Literature Search

Literature searches were completed in January 2021. Besides the database searches documented in the tables below, we searched a collection of government health agency, professional society, and medical center web sites that have been identified as evidence sources for CEP Rapid Guidance Summary reports relating to the SARS-CoV-2 pandemic. Because of the large number of sites that were searched, they are not tabulated here. Because of the rapid pace of COVID-19 research and the urgency of the topic, we searched the medRxiv preprint server for unpublished manuscripts (Table 5).

Table 1. Guideline searches

Database or organization	Keywords or syntax	Hits	Marked for retrieval	Included
ECRI Guidelines Trust	COVID or corona*	163	0	0

Table 2. Evidence clearinghouse searches

Search keywords	Evidence type	Hits	Marked for retrieval	Included
NICE Evidence Search (NHS)				
COVID and emergency	Guidance	†298	3	1
(limit to 2020-2021)	Systematic reviews	45	2	0
CEP Rapid Guidance Summary sources				
browsed		—	9	6

†–Only the 100 most relevant hits (as determined by NICE algorithm) were screened

Table 3. Medline search

Search	Syntax	Hits	Marked for retrieval	Included
1	(covid* or coronavirus or sars*).mp.	120,065	—	—
2	((emergency adj1 (room* or department* or care)) or short stay or observation).mp.	447,314	—	—
3	exp Emergency Service, Hospital/	81,858	—	—
4	CEP standard filter for clinical pathways and algorithms	738,504		
5	1 and (2 or 3) and 4	174	—	—
6	limit 5 to yr="2020 -Current"	166	—	—
7	(guideline* or guidance).mp. or exp Guideline/ or exp Practice Guideline/	605,437	—	—
8	1 and (2 or 3) and 7	254	—	—
9	limit 8 to yr="2020 -Current"	239	—	—

Search	Syntax	Hits	Marked for retrieval	Included
10	6 or 9	376	—	—
	delete 13 duplicate references within set	363	15	3

Table 4. EMBASE search

Search	Syntax	Hits	Marked for retrieval	Included
1	covid* OR coronavirus OR sars*	130,126	—	—
2	((emergency NEAR/1 (room* OR department* OR care)) OR short) AND stay OR observation	506,246	—	—
3	'emergency ward'/exp	155,888	—	—
4	CEP standard filter for clinical pathways and algorithms	1,067,048	—	—
5	#1 AND (#2 OR #3) AND #4	269	—	—
6	#5 AND [2020-2021]/py	232	—	—
7	'practice guideline'/de OR guideline:ti,ab OR guidance:ti,ab	646,943	—	—
8	#1 AND (#2 OR #3) AND #7	279	—	—
9	#8 AND [2020-2021]/py	253	—	—
10	#6 OR #9	451	—	—
	Delete 88 duplicate references	363	15	2

Table 5. medRxiv preprint search

Search	Syntax	Hits	Marked for retrieval	Included
1	"COVID* AND pathway*" and posted between "01 Jan, 2020 and 05 Feb, 2021"	67	-	—
2	"COVID* AND emergency AND algorithm*" and posted between "01 Jan, 2020 and 05 Feb, 2021"	†1,075	—	—
3	1 or 2		3	0

†-only the 150 most relevant hits (as determined by medRxiv algorithm) were screened

Results

Guidelines

We did not find any guidelines advocating a specific clinical algorithm or clinical prediction tool for patients presenting to the Emergency Department with suspected or confirmed COVID-19 disease. There is a disease severity algorithm that has been offered by the American College of Emergency Physicians, but its use is not referenced in an actual practice guideline.

Systematic reviews

We found no relevant systematic reviews in the peer-reviewed literature. An unpublished manuscript made available online on February 1 2021 (1) reports a systematic review of prediction models for severe illness and death from COVID-19. The authors found 46 articles meeting inclusion criteria (which were not limited to Emergency Department settings) and reported that nearly all of the studies were at relatively high risk of bias. No conclusions were drawn about comparative effectiveness of the models.

In their article introducing the 4C mortality score (4C), the ISARIC-4C consortium (2) included a table of reported mortality prediction scores and their diagnostic performance as measured by the area under the ROC curve (AUC[*]). We reproduce that table in Figure 1. All of the scores in the table have statistically significant predictive ability, but that does not mean that their predictive ability is clinically valuable. Uncertainty of the AUC for most scores is large enough that we cannot conclude that one score is superior to others, especially because scores were developed and tested on different groups of patients, raising the prospect that spectrum bias could confound comparisons. There is a non-peer-reviewed comparison of predictive tools on the PulmCrit web blog (3). The blogger (a critical care physician) notes problems with many published models including small derivation sets, absence of validation in an independent set of patients, and issues with reproducibility. He comments favorably on the 4C score, the quick COVID Severity Index (qCSI), and the NEWS2 score (which is not COVID-specific). Finally, it should be noted that some of the mortality prediction scores were developed using data from patients who had already been admitted to an inpatient hospital unit; they may not perform as well in the population of patients presenting to the Emergency Department.

[*] Area under the ROC is a popular metric for reporting the performance of a diagnostic test because it summarizes sensitivity and specificity in a single number and methods for calculating the AUC and its confidence interval are well documented, but it has two key weaknesses: the failure to account for the trade-off between sensitivity and specificity and the large degree to which AUC depends on performance of the test at clinically-irrelevant thresholds. (34)

Figure 1. Risk scores for predicting in-hospital mortality from COVID-19

Table 6 | Discriminatory performance of risk stratification scores within validation cohort (complete case) to predict in-hospital mortality in patients with covid-19

	Validation cohort*	
Model	No of patients with required parameters	AUROC (95% CI)
SOFA	197	0.614 (0.530 to 0.698)
qSOFA	19 361	0.622 (0.615 to 0.630)
Surgisphere†	18 986	0.630 (0.622 to 0.639)
SMARTCOP	486	0.645 (0.593 to 0.697)
NEWS	19 074	0.654 (0.645 to 0.662)
DL score†	16 345	0.669 (0.660 to 0.678)
SCAP	370	0.675 (0.620 to 0.729)
CRB65	19 361	0.683 (0.676 to 0.691)
COVID-GRAM†	1239	0.706 (0.675 to 0.736)
DS-CRB65	18 718	0.718 (0.710 to 0.725)
CURB65	15 560	0.720 (0.713 to 0.728)
Xie score†	1753	0.727 (0.701 to 0.753)
A-DROP	15 572	0.736 (0.728 to 0.744)
PSI	360	0.736 (0.683 to 0.790)
E-CURB65	1553	0.764 (0.740 to 0.788)
4C Mortality Score	14 398	0.774 (0.767 to 0.782)

AUROC=area under the receiver operating characteristic curve; covid-19=coronavirus disease 2019.
See appendix 13 for other metrics.
*Available data.
†Novel covid-19 risk stratification score.

Source: ISARIC-4C Consortium (2)

Primary literature

We found no studies assessing the performance of algorithms for managing patients presenting to the Emergency Department with suspected or confirmed COVID-19 disease. Several articles describe such algorithms (Table 8), but none of them reported on the effect on patient outcomes using those algorithms. Our search of non-peer-reviewed preprints also found no articles of this type.

There is patient outcome data for studies of risk stratification tools (Table 6). A wide variety of biomarkers and composite risk prediction tools have been reported in these studies, and there is no apparent consensus on which measurements are most appropriate. While some COVID-specific risk scores have been reported, few have been externally validated. Clinical prediction scores derived from internal validation alone limits reproducibility and generalizability. Other studies (4, 5) compared various laboratory test values in survivors and non-survivors; they may provide a starting point for developing a composite clinical prediction tool. Machine learning (6, 7) could potentially expedite clinical prediction tool development and improve performance, but validation of the tools is still necessary.

It should also be noted that most of these trials involved patients who became ill in the early weeks of the pandemic (i.e. quarters 1 and 2 of 2020). It can be expected that these patients will have had worse outcomes than patients currently presenting to Emergency Departments, since hospitals have had time to develop more effective treatments for COVID-19. We consider this worthy of a GRADE evidence downgrade on the basis of indirectness. Risk of bias in these studies is low (Table 7). Additional clinical prediction tools and tests of the predictive ability of lab test results have been reported as non-peer-reviewed and unpublished manuscripts. While some large validation cohorts have been used to demonstrate the predictive ability of these measures, the clinical impact of using them has still not been assessed, and our overall impression of this literature is that it is similar to the evidence that has been published to date.

Table 6. Primary studies reporting risk prediction tools for ED patients with COVID-19 disease

Author Location	Study design	Patients COVID status	Prediction tool (threshold)	Mortality	Outcome 2	Comment
Validated in independent patient sample						
Bauer (8) Germany	Cohort	Adult, ED N = 19 Confirmed	Calprotectin Lactate C-reactive protein Procalcitonin	Insufficient N for meaningful results (2 deaths)	Admission to ICU AUC 0.70 (0.42-0.99) AUC 0.80 (0.58-1.00) AUC 0.66 (0.36-0.96) AUC 0.60 (0.29-0.90)	Sensitivity and specificity not reported, (no threshold selected) but ROC curves shown. Published in letter form. Also reported multi-organ failure outcome.
Covino (9) Italy	Cohort	Age > 60, ED N = 210 Confirmed	COVID-GRAM (>17.7) qCSI (>5) ISARIC-4c (>8) NEWS (>4)	Sen. 88%, Spec. 61% Sen. 69%, Spec. 77% Sen. 88%, Spec. 56% Sen. 67%, Spec. 69%	Not reported	Patients intubated on arrival were excluded. Differences in diagnostic performance as measured with area under the ROC were not statistically significant.
ISARIC-4C (2) UK	Cohort	Adult, inpatient N = 57,824 Confirmed	4C score (>3) 4C score (>8) 4C score (>15)	Sen. 99.7%, Spec. 10.4% Sen. 92.5%, Spec. 38.6% Sen. 38.0%, Spec. 89.8%	Not reported	Mortality during index hospital stay. Additional thresholds reported in article. AUC 0.774 (95% CI 0.767-0.782)
King (10) USA	Cohort	Adult, inpatient or outpatient N = 13,323 Confirmed	VACO	AUC 0.84 (0.78-0.86)	Not reported	Sensitivity and specificity not reported. Index intended to provide an estimated percentage risk of mortality within 30 days. Patient group predominantly male (91%).

Author Location	Study design	Patients COVID status	Prediction tool (threshold)	Mortality	Outcome 2	Comment
Van Singer (11) Switzerland	Cohort	Adult, ED N = 76 Confirmed	 IL-6 sTREM-1 + resp. rate sTREM-1 alone Respiratory rate alone	(includes intubation) Sen. 100%, Spec. 22% Sen, 94%, Spec.61% Sen, 83%, Spec.81% Sen, 77%, Spec.76%	Oxygen required Sen. 98%, Spec. 50% NR NR NR	Not all patients had all tests. sTREM-1 alone had best performance for predicting mortality/intubation as measured by area under the ROC Area under the ROC for qSOFA score was not superior to sTREM-1.
Not validated in independent sample						
READY (6) USA	Cohort	Age not reported, ED or observation N = 197 Confirmed	 Machine learning algorithm MEWS (N = 183)	Not reported	Mechanical ventilation Sen. 90%, Spec., 58% Sen. 78%, Spec., 40%	Mechanical ventilation within 24 hours of arrival. threshold not reported. MEWS can be considered independently validated.
Russell (12) USA	Cohort	Adult, observation N = 116 Suspected	Composite tool (age > 48, bilateral infiltrates on CXR, O_2sat < 95%)	Not reported	Inpatient admission Sen., spec. not reported Diagnostic odds ratio 4.99	Composite tool of Hispanic ethnicity, bilateral infiltrates, O_2sat < 95) also performed well.

All studies published 2020

AUC–area under the ROC curve (95% confidence interval)

MEWS–Modified Early Warning Score

NEWS–National Early Warning Score

qCSI–Quick COVID-19 Severity Index

qSOFA– Quick Sepsis-related Organ Failure Assessment

sTREM-1– soluble triggering receptor expressed on myeloid cells

TREM-1–Triggering receptor expressed on myeloid cells

Table 7. Study quality appraisal: diagnostic/prognostic studies

Study	Bauer	Covino	ISARIC-4C	King	Van Singer	READY	Russell
1: Representative patient group	Y	Y	Y	N	Y	Y	Y
2. Reasonable gold standard	Y	Y	Y	Y	Y	Y	Y
3. Reference test given to all patients	Y	Y	Y	Y	Y	Y	Y
4. Same reference test for all patients	Y	Y	Y	Y	Y	Y	Y
5. Reference independent of study test	Y	Y	Y	Y	Y	N	N
6. Reference test blinded	Y	Y	Y	Y	Y	N	N
7. Study test blinded	Y	Y	Y	Y	Y	Y	Y
8. Avoided interpretation bias	Y	Y	Y	Y	Y	Y	Y
9. Reasonable attrition	Y	Y	Y	Y	Y	Y	Y
10. Funding source	N	Y	Y	Y	Y	Y	Y

CEP modified QUADAS scale available at www.uphs.upenn.edu/cep/methods

Table 8. Excluded studies

Study	Description	Reason for exclusion
Barie (13)	Cornell triage algorithm	No patient data.
Berdahl (14)	Cedars-Sinai discharge criteria	Reported outcomes only for discharged patients, no comparison group.
Bonetti (4)	Predictive value of lab test results	Reported only p-values, not sensitivity and specificity or AUC.
Casiraghi (15), Piva (16)	Bresica cohort study	Algorithm primarily for the purpose of deciding treatment.
Fan (17)	Comparison of existing severity scores for predicting mortality	Study performed in non-OECD country (China).
Manivel (18)	Adelaide ultrasound protocol	Use of imaging characteristics to characterize patient risk, no patient data.
Salinas (5)	Predictive value of lab test results	Reported only p-values, not sensitivity and specificity or AUC
Smargiassi (19)	Gemelli ultrasound algorithm	Use of imaging characteristics to characterize patient risk, no patient data.
Suh (20)	Columbia-Presbyterian clinical pathway	No patient data.
Varani (21, 22)	Monocyte distribution width as a predictive variable	Published as abstract only (note also these are redundant publications).
Wallace (23)	Birmingham triage algorithm	No patient data.
Yue (24)	Machine learning for evaluating CT scans	Study performed in non-OECD country (China).

All studies published 2020

Pathways and algorithms

Clinical pathways and discharge algorithms found by our searches are summarized in Table 9. We believe they are a representative sample of current practices in Emergency Department management of patients with suspected or confirmed COVID-19 disease. By the criteria of our pathway appraisal scale, they are all of low quality (Table 10). The urgent need for these pathways has precluded some of the steps necessary to developing an evidence-based and effective plan for managing these patients, including reviewing and citing evidence and documenting methods thoroughly. None of the pathways have been reported as an intervention or exposure in studies measuring clinical outcomes.

Some of the algorithms are framed as tools for determining the severity of disease and others for determining the most appropriate disposition, but the two purposes overlap: patients with mild disease who are at low risk of complications may be sent home for isolation and symptomatic treatment, patients at greater risks should be admitted to an inpatient unit, and patients whose illness is already severe should be admitted to a critical care unit.

Table 11 compares the criteria used by the different algorithms for determining when a patient is suitable for discharge from the Emergency Department. There are minor variations between different hospitals and health systems, but the general form of the pathways is similar. Oxygen saturation is included in all the algorithms; half of them specify that oxygen saturation should be measured after walking as well as at rest, to ensure that there is no desaturation with exertion. Respiratory rate is also included in the majority of algorithms. Blood pressure and heart rate are included in some but not all algorithms. Older patients are seen as greater risks for severe disease, but the algorithms have different approaches to this factor. Some of them use age as a specific criterion for keeping patients for observation, while others lump age in with other risk factors such as obesity, hypertension, chronic respiratory disease, and other chronic diseases.

Table 9. Clinical pathways and algorithms

Source	Recommendations	Comment
American College of Emergency Physicians (25)	Algorithm for estimating severity of disease in adult patients with suspected or confirmed COVID-19 disease.	Includes suggested disposition (discharge to home, observation, admit to inpatient unit, admit to ICU).
Brigham and Women's Hospital (26)	Algorithm for supporting admit/discharge decisions in adult patients with suspected or confirmed COVID-19 disease.	Not updated since April 2020. Part of a larger ED pathway.
Children's Hospital of Philadelphia (27)	Algorithm for estimating severity of disease in children with confirmed COVID-19 disease.	Includes suggested disposition (discharge with supportive care, admit to inpatient unit, admit to ICU).
University of Colorado (28)	Algorithm for supporting admit/discharge decisions in patients with suspected COVID-19 disease.	Part of a larger ED pathway, but does not address criteria for admitting to ICU.
NHS Scotland (29)	Algorithm for supporting admit/discharge decisions in adult patients with suspected COVID-19 disease.	

Source	Recommendations	Comment
NICE (UK) (30)	Algorithm for supporting admit/discharge decisions in patients with suspected COVID-19 disease.	Earlier algorithms from NHS hospitals are also available online.
Partners in Health (31)	Algorithm for supporting admit/discharge decisions in patients with confirmed COVID-19 disease.	
University of Washington (32)	Algorithm for estimating severity of disease in adult patients with suspected or confirmed COVID-19 disease.	No specific thresholds for vital signs other than SpO2 (≥ 94%).
World Health Organization (33)	Criteria for estimating severity of disease in adult patients or children with suspected COVID-19 disease.	Not in algorithm format.

All tools published in 2020

Table 10. Pathway appraisal

Pathway	ACEP	Brigham	CHOP	Colorado	NHS Scot.	NICE	PIH	Washington	WHO
1: Transparency of methods	C	C	B	C	C	C	B	B	A
2. Development group	B	C	C	C	C	C	B	C	A
3. Funding and conflict of interest	B	B	B	B	B	A	A	B	A
4. Evidence base	B	C	B	C	C	C	B	C	A
5. Patient population	B	B	B	B	A	B	B	B	B
6. Specific action criteria	A	B	C	B	B	B	B	B	B
7. Specific intervention details	C	C	C	C	C	C	C	C	C
8. Pathway validation	C	C	C	C	C	C	C	C	C
9. Pathway currency	B	B	B	B	B	B	B	B	B

Table 11. Algorithm criteria for discharge of patients from the Emergency Department

Element	ACEP	Brigham	Colorado	NHS Scot.	NICE	PIH	CHOP	Washington	WHO
Age	Risk factor	Risk factor	< 70	X	X	< 60	Algorithms did not include specific criteria.		
Respiratory rate	≤ 22	≤ 24	≤ 20	≤ 24	X	≤ 22			
SpO2 resting (room air)	≥ 93%	≥ 94%	≥ 91%	≥ 92%	> 94%	≥ 94%			
SpO2 walking (room air)	≥ 90% or <3% decrease	≥ 90%	X	X	< 3% decrease	X			
Heart rate	< 100	X	≤ 110	X	X	X			
Blood pressure	Normal for patient	X	X	X	X	≥ 90/60			
Immunocompromised	Not immunocomp.	Risk factor	Not immunocomp.	X	X	X			
Other					NEWS < 3				
Risk factors (and maximum number of risk factors before immediate discharge is not considered appropriate)	Age > 60, male, Black, BMI > 30, CAD, COPD, hypertension, diabetes, cancer, CVD (no more than 1)	Age > 65, asthma, COPD, hypertension, diabetes, CAD, CRD, liver disease, BMI > 40 (no more than 1)	Hypertension, CAD, COPD, diabetes, CRD (no more than 2)	Not specified	Not specified	Not specified			

All criteria must be met for discharge to be appropriate, unless otherwise noted.

Conclusion

While there are several published algorithms and clinical pathways for managing patients presenting to the Emergency Department with suspected or confirmed COVID-19 disease, none of them have been utilized in clinical trials to determine their effectiveness on patient outcomes. In an ideal study design, patients would be randomized to one pathway or another (or usual care), and outcomes such as mortality, readmission to the hospital, need for mechanical ventilation, and length of Emergency Department stay would be compared between the two groups.

Likewise, numerous patient characteristics, comorbidities, and laboratory test results have significant ability to predict adverse outcomes for COVID-19 patients (most studies looked at mortality as the outcome of interest). Evidence for this conclusion is of moderate strength based on AUC outcomes (Table 12). Composite risk prediction scores incorporating these measures have been developed, but again they have not been validated in comparative studies, so we are unable to determine which of these scores is most

effective. Furthermore, most of the assessments of these measures is commonly done using a one-dimensional statistic (area under the ROC) which does not take test threshold effects into account.

Additional weaknesses of the current evidence base are the small number of studies of any individual algorithm or predictive tool, the small number of adverse outcomes in some studies, and the reliance on studies from early in the pandemic, when effective treatments for the disease were lacking and effective critical care practices had not been identified.

Oxygen saturation is consistently used as a means for triaging patients in identified pathways and algorithms for managing COVID-19 patients presenting to the Emergency Department, though there is not agreement on the SpO_2 threshold that should be used (Table 11). Some centers measure SpO_2 in a walking test as well as when the patient is at rest. Respiratory rate is also widely but not unanimously used in these algorithms; use of other vital signs, comorbidities, and other risk factors varies.

Table 12. Evidence summary and GRADE analysis

Comparison	Outcome	Conclusion	Quantity and type of evidence	Starting level of evidence strength	Risk of bias	Inconsistency	Indirectness	Imprecision	Publication bias	Strong or very strong assn.	Dose-response	Confounders considered	Final level of evidence strength
Use of a clinical pathway or algorithm.	All clinical outcomes including mortality, readmission	No evidence											None
Use of a clinical risk prediction tool vs. no tool	All clinical outcomes including mortality, readmission	No evidence											None
Use of a clinical risk prediction tool vs. no tool	Predictive ability for mortality	Significant predictive value	4 diagnostic cohort studies	High	0	0	–1	0	0	0	0	0	Moderate
Use of one risk prediction tool vs. another tool	Predictive ability for mortality	Inconclusive	4 diagnostic cohort studies	High	0	0	–1	–1	0	0	0	0	Low

Additional clinical reviewers

- Keith C. Hemmert, MD (Emergency Medicine)
- Austin S. Kilaru, MD, MSHP (Emergency Medicine)

Conflict of interest disclosures

None of the authors have any relevant financial relationships with commercial interests associated with the subject of this review. The CEP conflict of interest disclosure policy is found at www.uphs.upenn.edu/cep/methods. CEP reports are funded by the University of Pennsylvania Health System.

References

1. Miller JL, Tada M, Goto M, Mohr N, Lee S. Prediction models for severe manifestations and mortality due to COVID-19: a rapid systematic review. medRxiv. 2021 January 01:2021.01.28.21250718. https://doi.org/10.1101/2021.01.28.21250718 (unpublished manuscript)

2. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, 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 September 09;370:m3339. https://doi.org/10.1136/bmj.m3339

3. PulmCrit – Oscar awards for the best COVID prognostic models [Internet].: EMCrit Project; 2020 [updated November 28; cited 6 Jan 2021]. Available from: https://emcrit.org/pulmcrit/covid-prognostic-models/.

4. Bonetti G, Manelli F, Patroni A, Bettinardi A, Borrelli G, Fiordalisi G, et al. Laboratory predictors of death from coronavirus disease 2019 (COVID-19) in the area of Valcamonica, Italy. Clin Chem Lab Med. 2020;58(7):1100-5. https://doi.org/10.1515/cclm-2020-0459

5. Salinas M, Blasco Á, Santo-Quiles A, Lopez-Garrigos M, Flores E, Leiva-Salinas C. Laboratory parameters in patients with COVID-19 on first emergency admission is different in non-survivors: Albumin and lactate dehydrogenase as risk factors. J Clin Pathol. 2020. http://dx.doi.org/10.1136/jclinpath-2020-206865

6. Burdick H, Lam C, Mataraso S, Siefkas A, Braden G, Dellinger RP, et al. Prediction of respiratory decompensation in COVID-19 patients using machine learning: the READY trial. Comput Biol Med. 2020;124. https://doi.org/10.1016/j.compbiomed.2020.103949

7. Bennett TD, Moffitt RA, Hajagos JG, Amor B, Anand A, Bissell MM, et al. The National COVID Cohort Collaborative: clinical characterization and early severity prediction. medRxiv. 2021 January 01:2021.01.12.21249511. (unpublished manuscript) https://doi.org/10.1101/2021.01.12.21249511

8. Bauer W, Diehl-Wiesenecker E, Ulke J, Galtung N, Havelka A, Hegel JK, et al. Outcome prediction by serum calprotectin in patients with COVID-19 in the emergency department. J Infect. 2020. https://doi.org/10.1016/j.jinf.2020.11.016

9. Covino M, De Matteis G, Burzo ML, Russo A, Forte E, Carnicelli A, et al. Predicting in-hospital mortality in COVID-19 older patients with specifically developed scores. J Am Geriatr Soc. 2021;69(1):37-43. https://doi.org/10.1111/jgs.16956

Penn Medicine

10. King JT, Yoon JS, Rentsch CT, Tate JP, Park LS, Kidwai-Khan F, et al. Development and validation of a 30-day mortality index based on pre-existing medical administrative data from 13,323 COVID-19 patients: The Veterans Health Administration COVID-19 (VACO) Index. PLoS ONE. 2020;15(11):e0241825. https://doi.org/10.1371/journal.pone.0241825

11. Van Singer M, Brahier T, Ngai M, Wright J, Weckman AM, Erice C, et al. COVID-19 risk stratification algorithms based on sTREM-1 and IL-6 in emergency department. J Allergy Clin Immunol. 2021;147(1):99-106.e4. https://doi.org/10.1016/j.jaci.2020.10.001

12. Russell FM, Wang A, Ehrman RR, Jacobs J, Croft A, Larsen C. Risk factors associated with hospital admission in COVID-19 patients initially admitted to an observation unit. Am J Emerg Med. 2020. https://doi.org/10.1016/j.ajem.2020.10.009

13. Barie PS, Bronstein M, Gibson CJ, Kelly AG, Lee C, Narayan M, et al. A simple three-tier classification system for triage, communication, and resource utilization by patients afflicted with COVID-19 disease. Surg Infect (Larchmt). 2020;21(8):726-7. https://doi.org/10.1089/sur.2020.146

14. Berdahl CT, Glennon NC, Henreid AJ, Torbati SS. The safety of home discharge for low-risk emergency department patients presenting with coronavirus-like symptoms during the COVID-19 pandemic: A retrospective cohort study. J Am Coll Emerg Physicians Open. 2020. https://doi.org/10.1002/emp2.12230

15. Casiraghi A, Domenicucci M, Cattaneo S, Maggini E, Albertini F, Avanzini S, et al. Operational strategies of a trauma hub in early coronavirus disease 2019 pandemic. Int Orthop. 2020;44(8):1511-8. https://doi.org/10.1007/s00264-020-04635-5

16. Piva S, Filippini M, Turla F, Cattaneo S, Margola A, De Fulviis S, et al. Clinical presentation and initial management critically ill patients with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection in Brescia, Italy. J Crit Care. 2020; 58:29-33. https://doi.org/10.1016/j.jcrc.2020.04.004

17. Fan G, Tu C, Zhou F, Liu Z, Wang Y, Song B, et al. Comparison of severity scores for COVID-19 patients with pneumonia: a retrospective study. Eur Respir J. 2020;56(3). https://dx.doi.org/10.1183%2F13993003.02113-2020

18. Manivel V, Lesnewski A, Shamim S, Carbonatto G, Govindan T. CLUE: COVID-19 lung ultrasound in emergency department. Emerg Med Australas. 2020;32(4):694-6. https://doi.org/10.1111/1742-6723.13546

19. Smargiassi A, Soldati G, Borghetti A, Scoppettuolo G, Tamburrini E, Testa AC, et al. Lung ultrasonography for early management of patients with respiratory symptoms during COVID-19 pandemic. J Ultrasound. 2020;23(4):449-56. https://dx.doi.org/10.1007%2Fs40477-020-00501-7

20. Suh EH, Bodnar DJ, Melville LD, Sharma M, Farmer BM. Crisis clinical pathway for COVID-19. Emerg Med J. 2020;37(11): 700-4. https://doi.org/10.1136/emermed-2020-209933

21. Varani M, Kodzhebash L, Rizkallah R, Ferrari P, Zambelli F, Castagnetti C, et al. Monocyte distribution width (MDW) as a new parameter useful in COVID-19 patients in emergency setting. Biochim Clin. 2020;44:S52. https://www.sibioc.it/bc/numero/bcnum/193

22. Varani M, Kodzhebash L, Rizkallah R, Ferrari P, Zambelli F, Castagnetti C, et al. MDW and IFCC laboratory parameters in COVID patients in Emergency setting. Biochim Clin. 2020;44:S52. https://www.sibioc.it/bc/numero/bcnum/193

23. Wallace DW, Burleson SL, Heimann MA, Crosby JC, Swanson J, Gibson CB, et al. An adapted emergency department triage algorithm for the COVID-19 pandemic. J Am Coll Emerg Physicians Open. 2020. https://doi.org/10.1002/emp2.12210

24. Yue H, Yu Q, Liu C, Huang Y, Jiang Z, Shao C, et al. Machine learning-based CT radiomics method for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. Ann Transl Med. 2020;8(14). https://doi.org/10.21037/atm-20-3026

25. COVID-19 Severity Classification Tool Now Available [Internet].: American College of Emergency Physicians; 2020 [updated August 17; cited 5 Jan 2021]. Available from: https://www.acep.org/corona/COVID-19-alert/covid-19-articles/covid-19-severity-classification-tool-now-available.

26. ED COVID Disposition Assessment Algorithm [Internet].: Brigham and Women's Hospital; 2020 [updated April 16; cited 6 Jan 2021]. Available from: https://bwh.covidprotocols.org/.

27. Clinical Pathway for Evaluation and Treatment of Patients with Active COVID-19 Infection [Internet].: Children's Hospital of Philadelphia; 2020 [updated November 23; cited 6 Jan 2020]. Available from: https://www.chop.edu/clinical-pathway/covid-disease-clinical-pathway.

28. COVID-19 initial ED evaluation and management [Internet].: University of Colorado Health; 2020 [updated November 23; cited 2 Feb 2021]. Available from: https://www.uchealth.org/today/clinical-practice-documents/.

29. COVID-19 Clinical Guidance for NHS Scotland: Emergency Department Management of Suspected COVID-19 in Adults Version 2. Scottish Government;NHS Scotland. 2020 April 16.

30. Reference Guide for Emergency Medicine [Internet]. London: National Institute for Health and Care Excellence; 2020 [updated November 01; cited 6 Jan 2021]. Available from: https://www.nice.org.uk/Media/Default/About/COVID-19/Specialty-guides/emergency-medicine-reference-guide.pdf.

31. PIH Algorithm for Initial Patient Assessment [Internet].: Partners in Health; 2020 [updated December 20; cited 6 Jan 2021]. Available from: https://covidprotocols.org/en/chapters/patient-assessment/#disease-severity-and-disposition-t4tualt31n26.

32. UW Medicine emergency department risk assessment algorithm for COVID-19 [Internet].: University of Washington; 2020 [updated November 09; cited 2 Feb 2021]. Available from: https://covid-19.uwmedicine.org/Pages/default.aspx.

33. World Health Organization. Clinical management of COVID-19: Interim guidance. Geneva: World Health Organization; 2020 May 27. https://www.who.int/publications-detail-redirect/clinical-management-of-covid-19

34. Mitchell MD. Validation of the summary ROC for diagnostic test meta-analysis: a Monte Carlo simulation. Acad Radiol. 2003;10(1):25-31. https://doi.org/10.1016/s1076-6332(03)80784-5

Appendix. CEP Trustworthy Pathway Appraisal Scale (development version)

Draft 4: December 24, 2019—Work in progress

The purpose of this appraisal tool is to identify characteristics of a clinical pathway or algorithm that bear on its reliability or applicability to particular patient groups. It is a pragmatic tool for appraisal of existing pathways, and not a comprehensive checklist of features that should be incorporated into a well-designed and evidence-based pathway. A grade of A, B, or C can be applied to each measure in the tool: A grades are characteristic of a higher quality pathway, while weaker pathways will be given C grades. The intermediate grade, B, is used when the evidence is less strong than for A but stronger than for C.

Knowing the strengths and weaknesses of a pathway will allow users to make more effective decisions about adaptation and use of existing clinical pathways. This appraisal tool is designed to identify specific types of weakness in a pathway.

1. Transparency of methods

	A	Methods for development of the pathway are thoroughly described.
	B	Methods for development of the pathway are partially described.
	C	Methods for development of the pathway are not described at all.

If the methods for pathway development are not described, we cannot determine the extent to which the pathway is evidence-based and unbiased. Criteria for answering this question have to be left imprecise since there is no one prescribed way to create a pathway. Consider key questions like how evidence was obtained and how the pathway was approved. If the pathway is described as based on a published guideline (or cites the published guideline), but no other development information is given, a grade of B should be assigned.

2. Development group

	A	Development group includes clinicians and staff from multiple specialties and multiple disciplines.
	B	Development group includes clinicians and staff from multiple specialties or multiple disciplines, but not both.
	C	Development group is from a single specialty or discipline or development group is not described.

Involvement of a diverse group of disciplines including physicians, nurses and other clinicians as appropriate) increases the value of a pathway because it increases our confidence that the pathway wasn't developed to further the interests of a narrow group of specialists.

3. Funding and conflict of interest

	A	Funding for the pathway is disclosed and no developers have a conflict of interest.
	B	No developers have a conflict of interest, but funding for the pathway is not disclosed.
	C	One or more developers has a significant potential conflict of interest or Both pathway and developer disclosures are not reported or funding for the pathway project has a significant potential COI

A conflict of interest brings the risk that a developer may promote products and services that further his or her own personal interests over products and services with stronger support in the evidence.

4. Evidence base

	A	Recommendations are supported by specific evidence, and the strength of that evidence is graded.
	B	Recommendations are supported by specific evidence, but are not graded.
	C	Evidence is not cited.

Our confidence in a pathway is increased if the developers can show that decisions are based on specific evidence. If the pathway states it is based on a specific guideline or guidelines, and the guideline grades the evidence, then an A grade may be assigned.

5. Patient population

	A	The pathway describes the inclusion/exclusion criteria for target population and includes variations for subgroups of interest.
	B	The pathway describes the inclusion/exclusion criteria target population, but not variations for subgroups of interest.
	C	The pathway does not describe the inclusion/exclusion criteria for target population.

The effectiveness or safety of a technology may differ for subgroups of patients such as elderly patients or patients with comorbidity. This question should be answered with an eye towards any obvious subgroups of patients we are likely to encounter in our system. Assign a B grade if the pathway has specific inclusion/exclusion criteria for patients, and assign an A grade if the pathway also includes alternate recommendations for different subgroups of patients.

in another stage should get a B grade. At minimum, in order to be a pathway in the first place, a set of recommendations must include an order in which actions should be taken and identify decision points where different actions may follow based on a test result or clinical response.

7. Specific action criteria

	A	The pathway consistently includes specific criteria or thresholds for taking actions.
	B	The pathway includes thresholds or criteria at some points.
	C	The pathway lacks specific action thresholds.

Just as important as the pathway of actions to take is the test threshold of other criteria for deciding whether or not to take an action. If action thresholds are lacking, implementation of the pathway is going to be less consistent.

8. Specific intervention details

	A	The pathway consistently includes specific and sufficient instructions to carry out the intervention (eg drug dosing)
	B	The pathway includes intervention details at some points.
	C	The pathway lacks specific intervention details.

Similarly, it is desirable for the pathway to include specific instructions for carrying out interventions. At minimum, a pathway needs to include actionable recommendations, using words like "should" and "should not." Details of routine aspects of patient care that do not relate to the core condition being addressed in the pathway are not necessary for a pathway to be graded A. For example, a pathway for postpartum care might not need details of how often vital signs should be measured, while a pathway for managing patients during alcohol withdrawal would.

9. The pathway has been validated

	A	The pathway has been validated outside the developing institution.
	B	The pathway has been validated, but only by the developers.
	C	No validation is reported.

We can have much more confidence in a pathway if there is evidence that use of the pathway actually improves patient outcomes. For purposes of this appraisal, validation can be any documented study that reports patient outcomes when the pathway is used.

10. The pathway is current

	A	The pathway is less than three years old and has an expiration or review date.
	B	The pathway is less than three years old but lacks an expiration or review date.
	C	The pathway is more than three years old.

If the pathway is old, there may be new evidence that would change the recommended practice. A pathway is more reliable if we know the developers understand that it needs periodic review to ensure that the evidence continues to support the pathway recommendations.

Reporting the results of this appraisal

As with other appraisal tools used by CEP, we do not attempt to convert the results of this appraisal into a numeric score. Quantitative scoring is necessarily grounded in a subjective judgment of the importance of individual components. Results are presented in a grid, so users can not only see the appraisal of each pathway, but also see where there are common areas of strength or weakness across all the available pathways. With that understanding, users can identify areas needing further investigation and review of the evidence.