
Article:
Dash, K., Goodacre, S. orcid.org/0000-0003-0803-8444 and Sutton, L. (2022) Composite outcomes in clinical prediction modeling: are we trying to predict apples and oranges? Annals of Emergency Medicine. ISSN 0196-0644

https://doi.org/10.1016/j.annemergmed.2022.01.046

Composite Outcomes in Clinical Prediction Modeling: Are We Trying to Predict Apples and Oranges?

Kieran Dash, MBChB, MPH*; Steve Goodacre, MSc, PhD; Laura Sutton, MSc, PhD

*Corresponding author. E-mail: kieranrdash@gmail.com.

Composite outcomes are widely used in clinical research. Existing literature has considered the pros and cons of composite outcomes in clinical trials, but their extensive use in clinical prediction has received much less attention. Clinical prediction assists decision-making by directing patients with higher risks of adverse outcomes toward interventions that provide the greatest benefits to those at the greatest risk. In this article, we summarize our existing understanding of the advantages and disadvantages of composite outcomes, consider how these relate to clinical prediction, and highlight the problem of key predictors having markedly different associations with individual components of the composite outcome. We suggest that a "composite outcome fallacy" may occur when a clinical prediction model is based on strong associations between key predictors and one component of a composite outcome (such as mortality) and used to direct patients toward intervention when these predictors actually have an inverse association with a more relevant component of the composite outcome (such as the use of a lifesaving intervention). We propose that clinical prediction scores using composite outcomes should report their accuracy for key components of the composite outcome and examine for inconsistencies among predictor variables. [Ann Emerg Med. 2022;■:1-8.]

0196-0644/$-see front matter

https://doi.org/10.1016/j.annemergmed.2022.01.046

INTRODUCTION

Health care research utilizes composite outcomes to capture the numerous potential benefits or risks of interventions.[1] A composite outcome is the combination of 2 or more outcomes, called component outcomes, into a singular outcome.[2,3] Composite outcomes often encapsulate broad concepts, with the common component outcomes in health care research including quality of care, patient-reported outcomes, morbidity, and mortality.[1,2] The composite outcome should be associated with the primary objective of a study and be both biologically plausible and meaningful to patients and clinicians.[4] In this article, we review the historical and intellectual groundwork on composite outcomes in research, draw the link to composite outcomes in clinical prediction scores, and then assess the advantages and disadvantages of using composite outcomes in clinical prediction.

ADVANTAGES AND DISADVANTAGES OF COMPOSITE OUTCOMES IN RESEARCH

The majority of the literature on composite outcomes focuses on their use in clinical trials. Composite outcomes are most commonly considered in trials of cardiovascular interventions.[5] Cardiovascular trials typically use the composite outcome of major adverse cardiac events, with combinations of coronary intervention, nonfatal myocardial infarction, and mortality used for the composite outcome.[6] In 2015, the European Network for Health Technology Assessment produced an updated guideline to provide a set of recommendations and aspects to be considered when assessing trials using composite outcomes.[3] The guideline was based on a literature review of the potential benefits and concerns of using composite outcomes in clinical trials, summarized in Table 1.

The main benefits include increased statistical efficiency, the potential to increase overall event rates when individual event rates are low, and the improved resource efficiency of trials.[3-5,7-9] Ross[4] argued that as population health generally improves and event rates decrease, composite outcomes would be needed to estimate event rates, without the need for very large samples. Additionally, composite outcomes allow researchers to assess the net clinical benefit of an intervention without having to choose a single outcome from many equally important outcomes of interest to clinicians and patients.[3,10]

However, the construction and reporting of composite outcomes have many challenges, which can lead to bias and difficulty interpreting the results. Cordoba et al,[5] in a

systematic review of clinical trials using composite outcomes, concluded that the construction of composite outcomes often lacks logic and is susceptible to post hoc choosing, or "cherry picking," of a favorable combinations of outcomes. Problems have been highlighted in the choice of component outcomes, but there is agreement that they should be of similar importance to patients and understandable to clinicians or end users.[2,4,9,10] Clinician-driven component outcomes, or "process outcomes," are of particular concern because they are subjective and prone to bias and may depend on individual clinician behavior, which makes the interpretation and generalization of results more challenging.[5,8] Ross[4] also argued that the reporting of the construction of composite outcomes and definitions of the component outcomes is generally inadequate.

The interpretation of composite outcomes in clinical trials can be further complicated by the assumption of uniform directionality.[11] This is the assumption that the intervention has a similar effect on each component of the composite outcome. Clinical trials can often report favorable results for composite outcomes when evidence of a favorable result for an important component outcome is lacking. For example, an intervention may reduce symptom burden but may not change mortality risk.[5,7,8] The assumption of uniform directionality can lead to errors in interpretation, which may lead to the underestimation or overestimation of the effect.[3] To address this issue, Ferreira-González et al[10] stated that trials should report individual component outcomes alongside the composite outcome. However, they also affirmed that reporting composite outcomes was appropriate to establish a potential net benefit.

Montori et al[12] proposed 3 questions to help clinicians decide whether to base clinical decisions on trials that have composite primary outcomes:

1. Are the component outcomes of similar importance to patients (recognizing that importance may be subjective and differ between patients)?
2. Do more- or less-important outcomes occur with similar frequencies?
3. Are the component outcomes likely to have similar risk reductions?

Finally, bias due to competing hazards may arise when one outcome precludes the occurrence of another, such as when death from other causes precludes the development of coronary heart disease. There is uncertainty in the current literature about the importance of this potential bias in composite outcomes.[9,10,13,14]

COMPOSITE OUTCOMES IN CLINICAL PREDICTION SCORES

Composite outcomes are often used in the development of clinical prediction models or scores. Clinical prediction scores are created to assist clinical decisionmaking and improve the quality of patient care.[15,16] Clinical prediction scores include diagnostic scores, which predict the probability of conditions being present; prognostic scores, which predict the probability of future outcomes; and prescriptive scores, which predict the probability of specific treatments or interventions being effective.[15] Clinical prediction scores help clinicians to estimate the risk of an adverse outcome and direct high-risk patients toward interventions that provide the greatest benefits to those at the greatest risk.[17] A clinical prediction score is usually composed of a small number of weighted predictor variables, created through logistic regression models or similar statistical techniques, to form a scoring system that stratifies patients' risks of outcomes.[18,19]

Table 1. Advantages and disadvantages of composite outcomes in clinical trials, adapted from the European Network of Health Technology Assessment.[3]

Advantages of Using Composite Outcomes	Disadvantages of Using Composite Outcomes
Statistical efficiency/reduce sample size requirements	Components are often unreasonably combined, inconsistently defined, and give opportunity for post hoc changes
Increased event rates	Improvement can be driven by a less-important component of composite outcome
Resource efficiency	Effects observed on separate components of a composite outcome may not be in same direction
Avoidance of arbitrary choice of single outcome when other outcomes may be equally important	Inclusion of unresponsive components may reduce the effect of the composite outcome
Avoiding adjustments for multiple comparisons	Clinician-driven outcomes may be prone to bias (eg, revascularization)
Estimates the net clinical benefit of intervention or treatment	Possible lack of relevance of component outcomes to patients
Effective treatment will be made available in timely manner	Alpha error must be adjusted to perform statistical inference on the components

Table 2. Examples of clinical prediction scores with composite outcomes used in EDs.

Clinical Prediction Score	Components of the Composite Outcome
HEART	• Acute myocardial infarction • PCI • CABG • Death
TIMI	• New or recurrent myocardial infarction • Severe recurrent ischemia requiring urgent revascularization • All-cause mortality
GRACE	• Myocardial infarction • Death
Pulmonary Embolus Severity Index	• 30-day mortality/inpatient mortality • Adverse events (including nonfatal cardiogenic shock or cardiorespiratory arrest)
Ottawa chronic obstructive pulmonary disease risk scale	• 30-day mortality • Admission to monitored unit/intubation/noninvasive ventilation • Myocardial infarction • Major procedure (CABG, PCI, other cardiac surgery, or new hemodialysis) • Return to ED for any related medical problem within 14 days
San Francisco Syncope Rule	• Death • Myocardial infarction/arrhythmia/pulmonary embolism/stroke/significant or subarachnoid hemorrhage • Return to ED and hospitalization for any related medical problem within 14 days
PRIEST COVID-19 clinical severity score	• 30-day mortality • Major organ support (respiratory, cardiovascular, or renal)

CABG, Coronary artery bypass; *GRACE*, Global Registry of Acute Coronary Event; *HEART*, History, Electrocardiogram, Age, Risk Factors and Troponin; *PCI*, percutaneous coronary intervention; *PRIEST*, Pandemic Respiratory Infection Emergency System Triage; *TIMI*, Thrombolysis in Myocardial Infarction.

Researchers have often used composite outcomes to develop prediction scores for use in the emergency department (ED). Examples are presented in Table 2 and include the History, Electrocardiogram, Age, Risk Factors and Troponin (HEART), Thrombolysis in Myocardial Infarction (TIMI), and Global Registry of Acute Coronary Event (GRACE) scores for patients presenting with chest pain, the Pulmonary Embolus Severity Index score, Ottawa chronic obstructive pulmonary disease risk scale, San Francisco Syncope Rule, and, recently, many coronavirus disease 2019 (COVID-19) severity scores.[20-25]

Clinical prediction scores are increasing in popularity, but systematic reviews have highlighted variable methodologic quality and reporting in underlying studies.[26-34] Cowley et al[18] reviewed the methodological standards and concluded that despite numerous updates to the methodologic standards first created in 1985 by Wasson et al,[35] there is little consensus on many aspects of the development of scores. Guidelines for the reporting of clinical prediction research and tools to assess the risk of bias are available; however, neither of these, nor current methodological standards, addresses the potential benefits or issues of composite outcomes in clinical prediction research.[36-38] Stiell and Wells[39] produced a review to help in the critical appraisal of the methodologic quality of clinical prediction scores for use in the ED but did not discuss the appraisal of composite outcomes. Jones and Platts-Mills[40] discussed the difficulty in the interpretation of composite outcomes in general studies for emergency clinicians, describing the composite outcome "clouding the clinical significance of component measures," but did not elaborate on the potential issues of composite outcomes in clinical prediction scores. Given the limited literature around composite outcomes in clinical prediction scores, we have drawn upon the literature relating to clinical trials and considered specific issues in clinical prediction to suggest the potential benefits and concerns in Table 3.

ADVANTAGES AND DISADVANTAGES OF COMPOSITE OUTCOMES IN CLINICAL PREDICTION

The major benefits of using composite outcomes in the development of clinical prediction scores are similar to those in clinical trials and include enhanced statistical power, increased outcome rates, and better resource

Table 3. Advantages and disadvantages of composite outcomes in clinical prediction.

Advantages of Using Composite Outcomes	Disadvantages of Using Composite Outcomes
Statistical efficiency/reduce sample size requirements	Components are often unreasonably combined, inconsistently defined, and give opportunity for post hoc changes
Increased event rates	Prediction can be driven by a less-important component of composite outcome*
Resource efficiency	Predictors may have different directions of association with separate components of a composite outcome*
Avoidance of arbitrary choice of single outcome when other outcomes may be equally important	Inclusion of unpredictable components may reduce the prediction of the model for the composite outcome*
Avoids the need for a separate model to predict each outcome*	Clinician-driven outcomes may be prone to bias (eg, revascularization)
	Possible lack of relevance of component outcomes to patients

*Highlights the advantages and disadvantages that are specific to clinical prediction.

efficiency. Composite outcomes also allow the development of a single model for predicting any undesirable outcome, rather than having to select a single outcome or develop a separate model for each outcome. Consequently, composite outcomes are appropriate when we intend to use the model to identify patients at low risk of any undesirable outcome who can, thus, avoid intervention. Prediction models based on composite outcomes are helpful when, as commonly used in the ED, they support discharge decisions for low-risk patients. Problems may arise, however, if we use the model to select patients for intervention rather than simply predict adverse outcomes.

Outcomes in clinical prediction can be broadly categorized into adverse events (such as death or nonfatal myocardial infarction) or interventions intended to prevent adverse events (such as coronary revascularization or critical care admission). Adverse events are usually objective and clearly undesirable, but health care intervention may prevent adverse events from occurring. If we use a clinical prediction model to select patients for intervention, we clearly need to predict cases in which intervention will prevent adverse events. Including interventions in a composite outcome can address this problem. However, interventions are "process outcomes" and have similar disadvantages to those identified in relation to clinical trials. "Process outcomes" may be dependent on clinical expertise or the availability of resources, such as access to ventilation support or the availability of revascularization procedures, making these outcomes less generalizable between health care systems. Clinicians may use clinical predictors to decide which patients receive the interventions, thus creating bias due to the lack of independence between predictors and outcomes. The independent adjudication of outcomes may reduce, but not eliminate, this risk of bias.

A further concern relates to uniform directionality, identified in relation to composite outcomes in clinical trials. Problems may arise in clinical prediction modeling if key predictors have markedly different associations with individual components of the composite outcome. We provide an example of this from the development of the Pandemic Respiratory Infection Emergency System Triage (PRIEST) COVID-19 clinical severity prediction score described in the supplementary material (available at http://www.annemergmed.com).[41] The PRIEST study has publications in *PLoS One* and *Emergency Medicine Journal* and *Resuscitation*.[42-46] This paper includes new analysis of the dataset that has not previously been published or presented.

The PRIEST score used a composite outcome that included the component outcomes of 30-day mortality and the receipt of major organ support. The score had a good prediction for the composite outcome in the validation cohort (c-statistic, 0.80), but the score provided better discrimination for predicting death without major organ support (c-statistic, 0.83) than for receipt of major organ support alone (c-statistic, 0.68).[41] To explore this concern, we created Least Absolute Shrinkage and Selection Operator (LASSO) multiple regression models to predict the individual components of the composite outcome. Table 4 compares the models for the composite outcome and the component outcomes and shows that several key predictor variables had markedly different associations with the component outcomes. Notably, lower baseline performance status and increasing age were positive predictors of mortality and the composite outcome but negative predictors of the receipt of major organ support.

This illustrates how using a composite outcome may lead to a model with lower predictive accuracy for a component outcome compared to a model developed specifically for that component. A more important practical consequence occurs if we use the model to select patients for intervention. The PRIEST severity score gives a higher

Table 4. Summary of logistic regression for component outcomes (mortality, any major organ support) and the PRIEST study composite outcome.

	30-day Mortality Unrestricted Regression Model	**Major Organ Support Unrestricted Regression Model**	**Composite Outcome Unrestricted Regression Model**
ROC area (95% CI)	0.86 (0.85-0.86)	0.78 (0.77-0.80)	0.82 (0.82-0.83)
Variable	**Adjusted Odds Ratio**	**Adjusted Odds Ratio**	**Adjusted Odds Ratio**
Individual Demographics			
Age (0=16-49 years, 1=50- 65 years, 2=66-80 years, 3=>80 years)	0=Reference 1=3.88 **(↑)** 2=8.28 **(↑)** 3=12.78 **(↑)**	0=Reference 1=1.74 **(↑)** 2=1.18 **(↑)** 3=0.59 **(↓)**	0=Reference 1=2.17 **(↑)** 2=2.58 **(↑)** 3=3.30 **(↑)**
Sex (0=female, 1=male)	0=Reference 1=1.36	0=Reference 1=1.36	0=Reference 1=1.38
Performance status (1=unrestricted normal activity, 2=limited strenuous activity, can do light activity, 3=limited activity, can self-care, 4=limited self-care, 5=bed/chair bound, no self-care)	1=Reference 2=1.16 **(↑)** 3=1.55 **(↑)** 4=2.22 **(↑)** 5=2.75 **(↑)**	1=Reference 2=0.88 **(↓)** 3=0.80 **(↓)** 4=0.59 **(↓)** 5=0.60 **(↓)**	1=Reference 2=0.96 **(↓)** 3=1.15 **(↑)** 4=1.50 **(↑)** 5=1.77 **(↑)**
Presenting Features			
Level of consciousness (1=alert, 2=verbal response, 3=pain response, 4=unresponsive)	1=Reference 2=2.22 **(↑)** 3=2.80 **(↑)** 4=4.29 **(↑)**	1=Reference 2=0.83 **(↓)** 3=0.66 **(↓)** 4=no value	1=Reference 2=1.98 **(↑)** 3=2.62 **(↑)** 4=4.28 **(↑)**
Respiratory Rate (NEWS score 0-3)	0=Reference 1=1.90 2=1.31 3=1.93	0=Reference 1=2.80 2=1.44 3=2.10	0=Reference 1=1.83 2=1.36 3=2.13
SaO_2 (NEWS score 0-3)	0=Reference 1=1.21 2=1.54 3=2.53	0=Reference 1=1.70 2=2.27 3=3.52	0=Reference 1=1.43 2=1.85 3=3.22
FiO_2 (NEWS score 0 or 2)	0=Reference 2=2.07	0=Reference 2=3.13	0=Reference 2=2.78
Pulse (NEWS score 0-3)	0=Reference 1=1.04 2=1.17 3=1.31	0=Reference 1=1.06 2=1.04 3=1.07	0=Reference 1=1.05 2=1.12 3=1.25
Systolic BP (NEWS score 0-3)	0=Reference 1=1.39 2=1.57 3=1.53	0=Reference 1=1.12 2=0.97 3=1.60	0=Reference 1=1.28 2=1.43 3=1.71
Temperature (NEWS score 0-3)	0=Reference 1=1.06 2=1.04 3=2.07	0=Reference 1=1.15 2=1.07 3=no value	0=Reference 1=1.14 2=1.09 3=1.89
Symptom duration (days)	1.00	No value	1.00
Shortness of breath	1.07	1.49	1.19
Cough	0.93	No value	0.99
Fever	0.91 (↓)	1.29 (↑)	1.04 (↑)
Respiratory distress	1.58	1.65	1.79
Respiratory exhaustion	1.06	1.79	1.72
Dehydration	1.47	No value	1.32

Table 4. Continued.

Variable	Adjusted Odds Ratio	Adjusted Odds Ratio	Adjusted Odds Ratio
Comorbidities			
Heart disease	1.06 (↑)	0.83 (↓)	0.95 (↓)
Hypertension	0.99 (↓)	1.14 (↑)	1.03 (↑)
Renal disease	1.23	1.32	1.35
Diabetes	1.09	1.29	1.27
Asthma	0.84	0.90	0.86
Other lung disease	0.86	0.69	0.78
Active malignancy (last 6 months)	1.86 (↑)	0.89 (↓)	1.52 (↑)
Steroid therapy	1.07	1.12	1.17
Immunosuppression	1.56	1.27	1.35
Medication count	0.99 (↓)	1.02 (↑)	0.99 (↓)

Adjusted odds ratio reported to 2 decimal places.
BP, Blood pressure; *CI*, confidence interval; *FiO^2*, fraction of inspired oxygen; *NEWS*, United Kingdom National Early Warning Score[50]; *ROC*, receiver operating characteristic; *SaO_2*, arterial oxygen saturation by pulse oximetry.

risk score for the composite outcome to an individual who is bed-bound and older than 80 years of age despite a negative association between these characteristics and the receipt of major organ support.[47] If we use the score in practice to select patients for organ support, we will select different patients from those selected for organ support in the derivation study and will probably select those less likely to benefit from organ support.

We could explain the variable associations between predictors and outcomes in terms of competing hazards. For example, patients with good performance status receive organ support and consequently survive, while those with poor performance status do not receive organ support and consequently die. If we assume that organ support prevents death, then these 2 components of the composite outcome are related, and it is reasonable to use the composite outcome to select patients for intervention. If, however, we anticipate that organ support will not prevent death, then it would not be reasonable to select patients for intervention using a model that was driven by the prediction of death.

We could describe this problem as a "composite outcome fallacy" in clinical prediction modeling. The fallacy occurs when we use the model to direct patients toward intervention despite key variables in the model predicting a lower probability of intervention in model derivation. It is a type of fallacy of composition, similar to the "gin and tonic fallacy," in which tonic is assumed to predict intoxication due to an observed association between gin and tonic and intoxication. The difference is that the "gin and tonic fallacy" relates to composite predictors rather than composite outcomes.

The "composite outcome fallacy" may occur when other prediction models in emergency medicine are used to select patients for intervention, notably clinical prediction models used for patients at risk of major adverse cardiac events—for example, the GRACE, TIMI, and HEART risk scores. A HEART score above the low-risk threshold (score $\geq$4) has better sensitivity for predicting mortality (95.0%) and myocardial infarction (97.6%) than coronary revascularization (89.7%), and the GRACE score has a higher discriminative value for predicting death (c-statistic, 0.81) than for the composite outcome of death and myocardial infarction (c-statistic, 0.73).[22,48] We theorize that this may be due to variables, such as increasing age and coronary risk factors in the HEART score, having positive or stronger associations with mortality than with revascularization interventions. A clinician may, therefore, make an error of interpretation if they assume that a higher HEART score indicates a greater need for intervention. Green and Schriger[49] recently appraised the methodology of the HEART score, but, to our knowledge, the specific issue we have highlighted has not been explored in the relevant literature.

SUGGESTIONS FOR FUTURE RESEARCH/ PRACTICE

Composite outcomes have important advantages in the development of clinical prediction scores, but clinicians need to be aware of their potential disadvantages, especially if they use the clinical prediction score to select patients for intervention. We have described a "composite outcome fallacy," whereby a clinical prediction score developed to predict a composite outcome could inappropriately direct patients toward intervention when key predictors are associated with a lower probability of intervention. We

suggest that clinical prediction scores using composite outcomes should report their accuracy for key components of the composite outcome and examine for inconsistencies among predictor variables. This would improve clinicians' understanding of the drivers of component outcomes within clinical prediction scores and assist appropriate use in decisionmaking. Current methodologic standards do not include the consideration of predictor variable association with the component outcomes in the appraisal of a clinical prediction score. We propose that methodologic standards for clinical prediction scores should include the examination of the association between key predictors and components of the composite outcome and the reporting of the accuracy of the scores for individual components of the composite outcome.

Supervising editor: David L. Schriger, MD, MPH. Specific detailed information about possible conflict of interest for individual editors is available at https://www.annemergmed.com/editors.

Author affiliations: From the School of Health and Related Research (ScHARR), University of Sheffield, Sheffield, United Kingdom (Dash, Goodacre, Sutton).

All authors attest to meeting the four ICMJE.org authorship criteria: (1) Substantial contributions to the conception or design of the work; or the acquisition, analysis, or interpretation of data for the work; AND (2) Drafting the work or revising it critically for important intellectual content; AND (3) Final approval of the version to be published; AND (4) Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Funding and support: By *Annals* policy, all authors are required to disclose any and all commercial, financial, and other relationships in any way related to the subject of this article as per ICMJE conflict of interest guidelines (see www.icmje.org). The authors received no specific funding to support the development or writing of this paper. The Pandemic Respiratory Infection Emergency System Triage study was funded by the United Kingdom National Institute for Health Research Health Technology Assessment program (project reference 11/46/07). The funder played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research, or the Department of Health and Social Care.

Publication dates: Received for publication December 18, 2021. Revision received January 19, 2022. Accepted for publication January 26, 2022.

REFERENCES

1. McKenna SP, Heaney A. Composite outcome measurement in clinical research: the triumph of illusion over reality? *J Med Econ*. 2020;23:1196-1204.
2. Peterson ED, DeLong ER, Masoudi FA, et al. ACCF/AHA 2010 position statement on composite measures for healthcare performance assessment: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures (Writing Committee to develop a position statement on composite measures). *Circulation*. 2010;121:1780-1791.
3. Endpoints used for Relative Effectiveness Assessment Composite Endpoints. European Network for Health Technology Assessment. Accessed November 10, 2021. https://www.eunethta.eu/wp-content/uploads/2018/03/composite_endpoints.pdf
4. Ross S. Composite outcomes in randomized clinical trials: arguments for and against. *Am J Obstet Gynecol*. 2007;196:119.e1-6.
5. Cordoba G, Schwartz L, Woloshin S, et al. Definition, reporting, and interpretation of composite outcomes in clinical trials: systematic review. *BMJ*. 2010;341:c3920.
6. Kip KE, Hollabaugh K, Marroquin OC, et al. The problem with composite end points in cardiovascular studies: the story of major adverse cardiac events and percutaneous coronary intervention. *J Am Coll Cardiol*. 2008;51:701-707.
7. Ferreira-González I, Busse JW, Heels-Ansdell D, et al. Problems with use of composite end points in cardiovascular trials: systematic review of randomised controlled trials. *BMJ*. 2007;334:786.
8. Freemantle N, Calvert M, Wood J, et al. Composite outcomes in randomized trials: greater precision but with greater uncertainty? *JAMA*. 2003;289:2554-2559.
9. McCoy CE. Understanding the use of composite endpoints in clinical trials. *West J Emerg Med*. 2018;19:631-634.
10. Ferreira-González I, Permanyer-Miralda G, Busse JW, et al. Methodologic discussions for using and interpreting composite endpoints are limited, but still identify major concerns. *J Clin Epidemiol*. 2007;60:651-657.
11. Armstrong PW, Westerhout CM. Composite end points in clinical research: a time for reappraisal. *Circulation*. 2017;135:2299-2307.
12. Montori VM, Permanyer-Miralda G, Ferreira-González I, et al. Validity of composite end points in clinical trials. *BMJ*. 2005;330:594-596.
13. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. *Circulation*. 2016;133:601-609.
14. Manja V, AlBashir S, Guyatt G. Criteria for use of composite end points for competing risks—a systematic survey of the literature with recommendations. *J Clin Epidemiol*. 2017;82:4-11.
15. Beattie P, Nelson R. Clinical prediction rules: what are they and what do they tell us? *Aust J Physiother*. 2006;52:157-163.
16. Laupacis A, Sekar N, Stiell IG. Clinical prediction rules: a review and suggested modifications of methodological standards. *JAMA*. 1997;277:488-494.
17. Hendriksen JM, Geersing GJ, Moons KG, et al. Diagnostic and prognostic prediction models. *J Thromb Haemost*. 2013;11(Suppl 1):129-141.
18. Cowley LE, Farewell DM, Maguire S, et al. Methodological standards for the development and evaluation of clinical prediction rules: a review of the literature. *Diagn Progn Res*. 2019;3:16.
19. Han K, Song K, Choi BW. How to develop, validate, and compare clinical prediction models involving radiological parameters: study design and statistical methods. *Korean J Radiol*. 2016;17:339-350.
20. Six AJ, Backus BE, Kelder JC. Chest pain in the emergency room: value of the HEART score. *Neth Heart J*. 2008;16:191-196.
21. Antman EM, Cohen M, Bernink PJ, et al. The TIMI risk score for unstable angina/non-ST elevation MI: a method for prognostication and therapeutic decision making. *JAMA*. 2000;284:835-842.
22. Fox KA, Dabbous OH, Goldberg RJ, et al. Prediction of risk of death and myocardial infarction in the six months after presentation with acute coronary syndrome: prospective multinational observational study (GRACE). *BMJ*. 2006;333:1091.
23. Aujesky D, Obrosky DS, Stone RA, et al. Derivation and validation of a prognostic model for pulmonary embolism. *Am J Respir Crit Care Med*. 2005;172:1041-1046.

24. Stiell IG, Perry JJ, Clement CM, et al. Clinical validation of a risk scale for serious outcomes among patients with chronic obstructive pulmonary disease managed in the emergency department. *CMAJ*. 2018;190:E1406-E1413.
25. Quinn JV, Stiell IG, McDermott DA, et al. Derivation of the San Francisco Syncope Rule to predict patients with short-term serious outcomes. *Ann Emerg Med*. 2004;43:224-232.
26. Bouwmeester W, Zuithoff NP, Mallett S, et al. Reporting and methods in clinical prediction research: a systematic review. *PLoS Med*. 2012;9:1-12.
27. Mallett S, Royston P, Dutton S, et al. Reporting methods in studies developing prognostic models in cancer: a review. *BMC Med*. 2010;8:20.
28. Collins GS, Mallett S, Omar O, et al. Developing risk prediction models for type 2 diabetes: a systematic review of methodology and reporting. *BMC Med*. 2011;9:103.
29. Collins GS, De Groot JA, Dutton S, et al. External validation of multivariable prediction models: a systematic review of methodological conduct and reporting. *BMC Med Res Methodol*. 2014;14:40.
30. Kleinrouweler CE, Cheong-See FM, Collins GS, et al. Prognostic models in obstetrics: available, but far from applicable. *Am J Obstet Gynecol*. 2016;214:79-90.e36.
31. Ettema RG, Peelen LM, Schuurmans MJ, et al. Prediction models for prolonged intensive care unit stay after cardiac surgery: systematic review and validation study. *Circulation*. 2010;122:682-689.
32. Collins GS, Omar O, Shanyinde M, et al. A systematic review finds prediction models for chronic kidney disease were poorly reported and often developed using inappropriate methods. *J Clin Epidemiol*. 2013;66:268-277.
33. Nayak S, Edwards DL, Saleh AA, et al. Performance of risk assessment instruments for predicting osteoporotic fracture risk: a systematic review. *Osteoporos Int*. 2014;25:23-49.
34. Altman DG. Prognostic models: a methodological framework and review of models for breast cancer. *Cancer Invest*. 2009;27:235-243.
35. Wasson JH, Sox HC, Neff RK, et al. Clinical prediction rules: applications and methodological standards. *N Engl J Med*. 1985;313:793-799.
36. Moons KG, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1-W73.
37. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170:51-58.
38. Green SM, Schriger DL, Yealy DM. Methodologic standards for interpreting clinical decision rules in emergency medicine: 2014 update. *Ann Emerg Med*. 2014;64:286-291.
39. Stiell IG, Wells GA. Methodologic standards for the development of clinical decision rules in emergency medicine. *Ann Emerg Med*. 1999;33:437-447.
40. Jones CW, Platts-Mills TF. Understanding commonly encountered limitations in clinical research: an emergency medicine resident's perspective. *Ann Emerg Med*. 2012;59:425-431.e11.
41. Goodacre S, Thomas B, Sutton L, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: the PRIEST observational cohort study. *PLoS One*. 2021;16: e0245840.
42. Goodacre S, Thomas B, Lee E, et al. Characterisation of 22445 patients attending UK emergency departments with suspected COVID-19 infection: observational cohort study. *PLoS One*. 2020;15: e0240206.
43. Goodacre S, Thomas B, Sutton L, et al. Derivation and validation of a clinical severity score for acutely ill adults with suspected COVID-19: the PRIEST observational cohort study. *P*LoS One. 2021;16: e0245840.
44. Thomas B, Goodacre S, Lee E, et al. Prognostic accuracy of emergency department triage tools for adults with suspected COVID-19: the PRIEST observational cohort study. *Emerg Med J*. 2021;38:587-593.
45. Sutton L, Goodacre S, Thomas B, et al. Do not attempt cardiopulmonary resuscitation (DNACPR) decisions in people admitted with suspected COVID-19: secondary analysis of the PRIEST observational cohort study. *Resuscitation*. 2021;164:130-138.
46. Goodacre S, Thomas B, Lee E, et al. Post-exertion oxygen saturation as a prognostic factor for adverse outcome in patients attending the emergency department with suspected COVID-19: a substudy of the PRIEST observational cohort study. *Emerg Med J*. 2021;38:88-93.
47. PRIEST COVID-19 clinical severity score. MDCalc. Accessed November 14, 2021. https://www.mdcalc.com/priest-covid-19-clinical-severity-score
48. Fernando SM, Tran A, Cheng W, et al. Prognostic accuracy of the HEART score for prediction of major adverse cardiac events in patients presenting with chest pain: a systematic review and meta-analysis. *Acad Emerg Med*. 2019;26:140-151.
49. Green SM, Schriger DL. A methodological appraisal of the HEART score and its variants. *Ann Emerg Med*. 2021;78:253-266.
50. Royal College of Physicians. National Early Warning Score (NEWS) 2: Standardising the assessment of acute-illness severity in the NHS **Updated report of a working party. 2017;17**.