Having a better understanding of the implications of ECMO on the lung allocation scoring and how it should be incorporated is important in trying to achieve the best possible outcome for those patients. Hayes and colleagues nicely demonstrate the proliferation in ECMO use over the years as well as the favorable survival across patients bridged with ECMO in high-volume lung transplant centers (6). This study highlights the need for more research surrounding the implications and role of ECMO in the care of patients undergoing lung transplant as well as further consideration of the most appropriate model for delivering high-quality, evidence-based care to high-risk patients undergoing lung transplant. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Laveena Munshi, M.D.
Interdepartmental Division of Critical Care Medicine
Mount Sinai Hospital
Toronto, Ontario, Canada

Eddy Fan, M.D., Ph.D.
Interdepartmental Division of Critical Care Medicine
and
Extracorporeal Life Support Program
Toronto General Hospital
Toronto, Ontario, Canada

References

1. Peek GJ, Mugford M, Tiruvoipati R, Wilson A, Allen E, Thalanany MM, Hibbert CL, Truesdale A, Clemens F, Cooper N, *et al*.; CESAR trial collaboration. Efficacy and economic assessment of conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure (CESAR): a multicentre randomised controlled trial. *Lancet* 2009;374:1351–1363.
2. Noah MA, Peek GJ, Finney SJ, Griffiths MJ, Harrison DA, Grieve R, Sadique MZ, Sekhon JS, McAuley DF, Firmin RK, *et al*. Referral to an extracorporeal membrane oxygenation center and mortality among patients with severe 2009 influenza A(H1N1). *JAMA* 2011;306: 1659–1668.
3. Tudorache I, Sommer W, Kühn C, Wiesner O, Hadem J, Fühner T, Ius F, Avsar M, Schwerk N, Böthig D, *et al*. Lung transplantation for severe pulmonary hypertension: awake extracorporeal membrane oxygenation for postoperative left ventricular remodelling. *Transplantation* 2015;99:451–458.
4. Cypel M, Keshavjee S. Extracorporeal life support as a bridge to lung transplantation. *Clin Chest Med* 2011;32:245–251.
5. Mohite PN, Sabashnikov A, Reed A, Saez DG, Patil NP, Popov AF, DeRobertis F, Bahrami T, Amrani M, Carby M, *et al*. Extracorporeal life support in "awake" patients as a bridge to lung transplant. *Thorac Cardiovasc Surg* 2015;63:699–705.
6. Hayes D Jr, Tobias JD, Tumin D. Center volume and extracorporeal membrane oxygenation support at lung transplantation in the lung allocation score era. *Am J Respir Crit Care Med* 2016;194:317–326.
7. Barbaro RP, Odetola FO, Kidwell KM, Paden ML, Bartlett RH, Davis MM, Annich GM. Association of hospital-level volume of extracorporeal membrane oxygenation cases and mortality: analysis of the extracorporeal life support organization registry. *Am J Respir Crit Care Med* 2015;191:894–901.
8. Hayanga JA, Lira A, Vlahu T, D'Cunha J, Hayanga HK, Girgis R, Aboagye J, Khaghani A. Procedural volume and survival after lung transplantation in the United States: the need to look beyond volume in the establishment of quality metrics. *Am J Surg* 2016;21: 671–676.
9. Hayanga JA, Lira A, Vlahu T, Yang J, Aboagye JK, Hayanga HK, Luketich JD, D'Cunha J. Lung transplantation in patients with high lung allocation scores in the US: evidence for the need to evaluate score specific outcomes. *J Transplant* 2015;2015:836751.
10. Shafii AE, Mason DP, Brown CR, Thuita L, Murthy SC, Budev MM, Pettersson GB, Blackstone EH. Too high for transplantation? Single-center analysis of the lung allocation score. *Ann Thorac Surg* 2014;98:1730–1736.
11. Hoopes CW, Kukreja J, Golden J, Davenport DL, Diaz-Guzman E, Zwischenberger JB. Extracorporeal membrane oxygenation as a bridge to pulmonary transplantation. *J Thorac Cardiovasc Surg* 2013; 145:862–867. [Discussion, pp. 867–868.]
12. Kahn JM, Goss CH, Heagerty PJ, Kramer AA, O'Brien CR, Rubenfeld GD. Hospital volume and the outcomes of mechanical ventilation. *N Engl J Med*. 2006 355:41–50.

The Misapplication of Severity-of-Illness Scores toward Clinical Decision Making

Decision tools and severity-of-illness scores are two distinct entities with different derivation methodologies and applications. In real-world practice, however, decision tools and severity-of-illness scores are frequently used interchangeably, resulting in the misapplication of severity-of-illness scoring systems to inform clinical decision making. The Sepsis III clinical criteria for the assessment of sepsis (in particular, the quick Sequential Organ Failure Assessment [qSOFA] score) and associated recommendations, recently presented by the Society of Critical Care Medicine and the European Society of Intensive Care Medicine Sepsis III task force (1, 2), may represent another such misapplication. As we discuss herein, clinicians and investigators must be cognizant of the distinction between severity-of-illness scores and decision tools to improve clinical decision making and ultimately prevent patient harm.

Severity-of-Illness Scores and Decision Tools

Severity-of-illness scores estimate the probability of an outcome of interest (e.g., mortality) on the basis of known patient characteristics. Severity-of-illness scores can be useful in estimating baseline risk in observational studies, performing hospital-to-hospital adjusted outcome comparisons, isolating specific risk strata for investigation, and providing prognostic information to patients and families. Severity-of-illness scores are generally derived through the application of stepwise logistic regression to an observational cohort; the resulting model is then judged based on overall model discrimination (often the area under the receiver operating characteristics curve [AUROC]), and calibration is then validated in a separate cohort. As such, severity-of-illness scores often seek the model with the best possible combination

of sensitivity and specificity (i.e., AUROC) across various cut-off values. In contrast, a decision tool must account for the risks/benefits of the actions recommended by the model, often balancing one desired outcome at the expense of another. The desired test characteristics of decision tools may vary based on the relative consequences of false positives/negatives. A clinician deciding whether to test for a highly morbid disease process, for instance, may prefer a decision tool with high sensitivity as opposed to one with an excellent overall AUROC.

As one example, the CURB-65 (confusion, uremia, elevated respiratory rate, hypotension, and age >65) score is a severity-of-illness score devised to predict mortality in patients with community-acquired pneumonia (CAP). In a derivation dataset containing 718 patients admitted to the hospital with CAP, the CURB-65 score effectively stratified patients by increasing risk of mortality: score 0, 0.7%; score 1, 3.2%; score 2, 3%; score 3, 17%; score 4, 41.5%; and score 5, 57% (3).

Moving beyond risk stratification, however, the authors of the CURB-65 study suggest that patients with the lowest predicted mortality (score 0 or 1) can likely be discharged to home, whereas the patients with scores predicting an intermediate risk of mortality (score 2 or 3) or high risk of mortality (score 4 or 5) should be assessed for hospitalization and admission to an intensive care unit (ICU), respectively (3). This recommendation has since made its way into national guidelines for the management of CAP, including the British Thoracic Society guidelines (4). However, the CURB-65 model predicts risk of inpatient mortality as the endpoint as opposed to an actionable decision. In other words, a given CURB-65 score provides a risk of mortality but does not evaluate safe discharge or discharge without need for readmission and/or outpatient death. The CURB-65 score also does not allow for assessment of the mortality rate modified by inpatient care (e.g., high-flow supplemental oxygen, mechanical ventilation, and vasopressors). In addition, other components of decision making, such as cost savings achieved by discharging patients home, may be eclipsed by the cost of subsequent admissions, even if only in a subset of patients.

The Canadian C-Spine Rule, in contrast, was originally conceived as a decision tool. In the derivation of the Canadian C-Spine Rule, a prespecified list of 25 clinical characteristics believed to be predictive of cervical spine injury were prospectively recorded in patients with blunt head/neck trauma. Given the potential for devastating injury after missed cervical spine injury, the authors *a priori* determined a desired sensitivity of 100% (95% confidence interval, 97–100%) for the final decision tool. Through recursive partitioning, the authors developed a rule that boasted 100% sensitivity (95% confidence interval, 98–100%) for detecting clinically important cervical spine injuries. Furthermore, the authors predicted a reduction in cervical spine radiography by 15.5% (5). The Canadian C-Spine Rule has since been prospectively validated and compared with other existing C-spine rules (e.g., the NEXUS [National Emergency X-Radiography Utilization Study] tool) (6) and evaluated in a randomized trial (7). The robust results of these studies provide clinicians with an approach for evaluating patients with potential cervical spine injury that allows them to safely forgo imaging in some cases. Clearly, a severity-of-illness score (especially one using mortality as an endpoint) would not capture the full complexity of medical decision making as it relates to C-spine injuries.

Sepsis III Definitions and Operational Use of the qSOFA Score

The Sepsis III definition of sepsis is "life-threatening organ dysfunction due to a dysregulated host response to infection" (1). In an attempt to operationalize the definition of sepsis, a number of severity-of-illness scoring systems that tie worsening organ dysfunction to risk of mortality were evaluated. Through secondary analysis of the electronic health records of more than 70,000 non-ICU patients at the University of Pittsburgh Medical Center, the authors determined the test characteristics for mortality prediction of the systemic inflammatory response syndrome criteria (AUROC, 0.76), SOFA score (AUROC, 0.79), Logistic Organ Dysfunction System (AUROC, 0.81), and the newly derived qSOFA (AUROC, 0.81) (1).

Although these results arguably demonstrate that qSOFA is a valid severity-of-illness score for predicting mortality in patients with suspected infection outside of the ICU, the use of these findings for decision making is problematic. First, as part of the data-driven derivation of the qSOFA score, investigators selected the worst values ranging from 48 hours before until 24 hours after the "onset" of infection. This methodology, although important for reducing missing data in a retrospective cohort, limits the application of the score for clinicians making decisions at the bedside. An emergency department clinician encountering a normotensive but confused patient with pneumonia (qSOFA = 1) does not have the luxury of knowing whether that patient will develop hypotension at some point in the following 24 hours. If that patient does progress and ultimately meets criteria for the Sepsis III definition of "sepsis," they may have already suffered from delayed intensification of care (e.g., narrower initial antibiotics, incomplete testing for occult organ dysfunction) and inappropriate triage decisions (e.g., home or ward vs. ICU). CURB-65 suffered a similar limitation by evaluating admitted patients undergoing inpatient therapy to determine a rule for discharge. In contrast, the Canadian C-Spine rule chose a patient population in which all were being considered for evaluation of a C-spine injury at the time of entry into the analysis

The current recommendations encourage the use of qSOFA as a screening test for sepsis and suggests that the "qSOFA criteria be used to prompt clinicians to further search for organ dysfunction, to initiate or escalate therapy as appropriate, and to consider referral to critical care or increased monitoring" (2). As highlighted above, however, applying a severity-of-illness score to derive a decision rule matches the wrong statistical approach with a clinical application. What specific actions, for instance, should be taken by providers when patients meet two of three qSOFA criteria? What are the complication rates and financial costs associated with those actions? What is the mortality rate of patients with infection and occult organ dysfunction missed by the qSOFA criteria? At present, the current recommendations do not address these questions, and the proposed operational flow diagram is not matched with the analyses that were performed. In contrast, when

the Canadian C-Spine criteria are met, providers know that C-spine imaging does not need to be performed—that is, an actionable endpoint on the right population has been assessed with proper methodology yielding excellent sensitivity and negative predictive value for that action.

Conclusions and Future Directions

Severity-of-illness scores and decision-support tools are two distinct entities. Whereas decision tools are developed with specific decisions and evaluation of their consequences in mind, severity-of-illness scores are more applicable for population-based prognostication and identification of similar-risk cohorts. For these reasons, we caution against interchanging severity-of-illness scores—including CURB-65 and qSOFA—with decision tools. Furthermore, we suggest that the principles of decision analysis be applied to deriving guidelines for care of patients with infection. ■

Author disclosures are available with the text of this article at www.atsjournals.org.

Ari Moskowitz, M.D.
Department of Medicine
Massachusetts General Hospital
Boston, Massachusetts

Lars W. Andersen, M.D., M.P.H.
Research Center for Emergency Medicine
Aarhus University Hospital
Aarhus, Denmark

Michael Cocchi, M.D.
Department of Anesthesia Critical Care
Beth Israel Deaconess Medical Center
Boston, Massachusetts

Michael W. Donnino, M.D.
Department of Emergency Medicine
Beth Israel Deaconess Medical Center
Boston, Massachusetts

References

1. Seymour CW, Liu VX, Iwashyna TJ, Brunkhorst FM, Rea TD, Scherag A, Rubenfeld G, Kahn JM, Shankar-Hari M, Singer M, *et al*. Assessment of clinical criteria for sepsis: for the Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:762–774.
2. Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche JD, Coopersmith CM, *et al*. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315:801–810.
3. Lim WS, van der Eerden MM, Laing R, Boersma WG, Karalus N, Town GI, Lewis SA, Macfarlane JT. Defining community acquired pneumonia severity on presentation to hospital: an international derivation and validation study. *Thorax* 2003;58:377–382.
4. Lim WS, Baudouin SV, George RC, Hill AT, Jamieson C, Le Jeune I, Macfarlane JT, Read RC, Roberts HJ, Levy ML, *et al*.; Pneumonia Guidelines Committee of the BTS Standards of Care Committee. BTS guidelines for the management of community acquired pneumonia in adults: update 2009. *Thorax* 2009;64:iii1–iii55.
5. Stiell IG, Wells GA, Vandemheen KL, Clement CM, Lesiuk H, De Maio VJ, Laupacis A, Schull M, McKnight RD, Verbeek R, *et al*. The Canadian C-spine rule for radiography in alert and stable trauma patients. *JAMA* 2001;286:1841–1848.
6. Stiell IG, Clement CM, McKnight RD, Brison R, Schull MJ, Rowe BH, Worthington JR, Eisenhauer MA, Cass D, Greenberg G, *et al*. The Canadian C-spine rule versus the NEXUS low-risk criteria in patients with trauma. *N Engl J Med* 2003;349:2510–2518.
7. Stiell IG, Clement CM, Grimshaw J, Brison RJ, Rowe BH, Schull MJ, Lee JS, Brehaut J, McKnight RD, Eisenhauer MA, *et al*. Implementation of the Canadian C-Spine Rule: prospective 12 centre cluster randomised trial. *BMJ* 2009;339:b4146.