Measuring Clinical Performance: Make Sure the Ruler Is Accurate*

Neil MacIntyre, MD
Division of Pulmonary and Critical Care Medicine
Duke University Medical Center, Duke University
Durham, NC

Measuring clinical performance is an essential component of assuring high-quality care. Accessing large databases to accomplish this is a methodology used by many investigators and is a tool becoming increasingly popular among payers and regulators. In accessing large databases, however, one has to be careful to assure that what is being measured accurately reflects clinical realities. Put another way, what may be easy to measure in a database may not be a true reflection of clinical performance. In this issue of *Critical Care Medicine*, Kramer et al (1) from the Cerner Corporation (the developers of the popular Acute Physiology and Chronic Health Evaluation [APACHE] system) report a methodology to assess clinical performance in the management of mechanical ventilation (MV). Although a laudable goal, I have serious concerns that their measured parameters may not be a reliable indicator of clinical quality.

In this Cerner study, over 80,000 patients in 48 U.S. hospitals using the APACHE system are evaluated. In the first 33,217 patients, a sophisticated model using APACHE data is used to construct a prediction equation for the duration of MV. This equation is then validated in a subsequent cohort of 23,065 patients. The investigators argue that the results demonstrate that this model, while not very useful in predicting an individual patient outcome, is a robust tool to assess performance of ICUs and health systems.

However, the "devil is in the details" and I have serious concerns about the way the critical endpoint, duration of MV was quantified. The Cerner investigators defined duration of MV as beginning with the institution of MV and ending when 1) MV was not required for 2 hours, 2) transfer to another facility, or 3) death. Defining duration of MV, however, is much more complex than that.

First, the 2-hour duration used by the Cerner investigators is shorter than virtually all clinicians would accept as reflective of permanent ventilator liberation. Indeed, 2 hours is the recommended duration of a spontaneous breathing trial (SBT) to assess the possibility of ventilator liberation (2). Clinically, patients tolerating (passing) the SBT are then assessed for the continued need of an artificial airway and decisions are made to extubate. They are then generally intensively monitored for the next 24–48 hours to assure clinical stability. With this approach, reintubations and reinstitution of MV still occur and the prevalence is reported to be as high as 20–25% of patients (3–6). Indeed, some have argued that a nominal reintubation rate (10–15%) assures a desirable balance between liberation aggressiveness and clinical restraint (6).

Defining liberation is made more complex in the setting of tracheostomies where patients can experience many hours of trach collar breathing but still need the MV for part of the day or night. Finally, the increasing use of noninvasive ventilation post extubation has further clouded the definitions of ventilator dependence/liberation. The Cerner 2-hour definition of no MV would seem a better reflection of time to a successful SBT (in itself a potential interesting clinical indicator), not a reliable or accurate reflection of the total duration of MV in many patients.

A second concern is the use of death or transfer as a marker of the end of the need for MV. This occurred in 16.4% (development set) to 20.7% (validation set) of patients. Although both of these clearly mark the end of the need for MV in that institution, neither makes any clinical sense as a marker for successful ventilator management. Although it could be argued that eliminating these subjects from the analysis might benefit ICUs with high death rates or rapid transfer policies (the most challenging patients are eliminated early), equating them to successfully liberated patients in terms of MV duration is counterintuitive.

Taken together, these issues should raise concern that this proposed model will be a fair tool to judge clinicians, ICUs, or health systems. Again, although the goals of this analysis are laudable, the model in its present form seems to be a fundamentally flawed ruler.

REFERENCES

1. Kramer AA, Gershengorn HB, Wunsch H, et al: Variations in Case-Mix–Adjusted Duration of Mechanical Ventilation Among ICUs. *Crit Care Med* 2016; 44:1042–1048
2. ACCP/SCCM/AARC Task Force: Evidence based guidelines for weaning and discontinuing mechanical ventilation. *Chest* 2001; 120(6 Suppl):375S–95S
3. Miu T, Joffe AM, Yanez ND, et al: Predictors of reintubation in critically ill patients. *Respir Care* 2014; 59:178–185
4. Liu Y, Wei LQ, Li GQ, et al: A decision-tree model for predicting extubation outcome in elderly patients after a successful spontaneous breathing trial. *Anesth Analg* 2010; 111:1211–1218
5. Menon N, Joffe AM, Deem S, Yanez MD, et al: Occurrence and complications of tracheal reintubation in critically ill adults. *Resp Care* 2012;57:1555–63
6. Kapnadak SG, Herndon SE, Burns SM, et al: Clinical outcomes associated with high, intermediate, and low rates of failed extubation in an intensive care unit. *J Crit Care* 2015; 30:449–54

***See also p. 1042.**

Key Words: clinical quality metrics; clinical outcome modelling; liberation from mechanical ventilation; mechanical ventilation

The author has disclosed that he does not have any potential conflicts of interest.

DOI: 10.1097/CCM.0000000000001682