SAPS3-CNIV score to predict hospital mortality following noninvasive ventilation: methodology insights

To the Editor:

We have read with great interest and congratulate HUSEINI *et al.* [1] on the external validation of the SAPS3-CNIV score (Simplified Acute Physiology Score 3 customised for noninvasive ventilation) to predict hospital mortality. HUSEINI *et al.* [1] concluded that SAPS3-CNIV did not improve prediction of mortality in patients over SAPS3. This scoring system included the SAPS3, haemoglobin, carbon dioxide tension, lactate, do not resuscitate (DNR) orders and aetiology of respiratory failure. However, we consider that there are some key aspects that need to be taken into account for a proper clinical extrapolation.

In routine clinical work, for the prediction of hospital mortality, it is important to use a simple scoring system [2]. This helps to make decisions and improve clinical outcomes. However, in our opinion, a scoring system may be developed more rigorously. From a methodology point of view, in the development of SAPS3-CNIV by MARTINEZ-URBISTONDO *et al.* [3], the following three points should be noted.

First, they only used 241 patients to derive this score and did not calculate the sample size. The in-hospital mortality was 32.4% [3]. However, if the power is to be adequate to reach 75% sensitivity and 75% specificity, with $\alpha=0.05$ and maximum marginal error of estimate of 5%, the smallest sample size is 890 [4]. Thus, we suggest that the statistical power needs to be re-evaluated for a clinical extrapolation.

Secondly, they did not specify when the variables were collected, such as at the beginning of noninvasive ventilation or after 1 h of treatment.

Thirdly, the variables in SAPS3-CNIV were dichotomised. The predictive power is not different in patients with lactate of 3 or 15 $mg \cdot dL^{-1}$, as the variable was dichotomised as >2 or $\leqslant 2$ $mg \cdot dL^{-1}$. The same problems were present in other variables. Obviously, this is not consistent with reality and suggests that these variables should be classified according to their gradients and categorised by various measured points.

Furthermore, two practical questions are important to add to these comments. 1) When was the DNR order made? 2) When was the patient admitted: to the intensive care unit or just before death? This has implication that, when measured at different time points, the scoring system may have different predictive power.

These limitations could explain why the predictive power is low in the validation study [1]. Further studies focused on this topic should be well designed to develop a general scoring system to guide clinical work.

Jun Duan[1], Bushra Mina[2] and Antonio M. Esquinas[3]
[1]Dept of Respiratory and Critical Care Medicine, the First Affiliated Hospital of Chongqing Medical University, Chongqing, China. [2]Dept of Medicine, Division of Pulmonary and Critical Care Medicine, Northwell Health, Lenox Hill Hospital, New York, NY, USA. [3]Intensive Care and Non-invasive Ventilatory Unit, Hospital Morales Meseguer, Murcia, Spain.

Correspondence: Jun Duan, Dept of Respiratory and Critical Care Medicine, the First Affiliated Hospital of Chongqing Medical University, Chongqing, 400016, China. E-mail: duanjun412589@163.com

Conflict of interest: None declared.

References

1 Huseini T, Manners D, Jones S, *et al.* External validation of the SAPS3-CNIV score to predict hospital mortality following noninvasive ventilation: a retrospective single-centre study. *ERJ Open Res* 2019; 5: 00232-2018.

2 Breslow MJ, Badawi O. Severity scoring in the critically ill: part 1 – interpretation and accuracy of outcome prediction scoring systems. *Chest* 2012; 141: 245–252.

3 Martinez-Urbistondo D, Alegre F, Carmona-Torre F, *et al.* Mortality prediction in patients undergoing non-invasive ventilation in intermediate care. *PLoS One* 2015; 10: e0139702.

4 Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med* 1996; 3: 895–900.