European Journal of Radiology 120 (2019) 108663

Contents lists available at ScienceDirect

European Journal of Radiology

journal homepage: www.elsevier.com/locate/ejrad

Lung ultrasound in evaluating the severity of neonatal respiratory distress syndrome: Methodological issues on diagnostic value and prediction to avoid misinterpretation

ARTICLE INFO

Keywords:
Diagnostic accuracy
Lung ultrasound
Mechanical ventilation
Neonatal respiratory distress syndrome
Prediction

Dear Editor,

I was interested to read the article titled "Diagnostic value of lung ultrasound in evaluating the severity of neonatal respiratory distress syndrome" by Pang H and colleagues [1] in the Jul 2019 issue of the Eur J Radiol. It is still unclear whether lung ultrasound (LUS) can be used to evaluate the severity of neonatal respiratory distress syndrome (NRDS). The purpose of this study was to evaluate the role of LUS in assessing NRDS. They reported that the receiver operative curve (ROC) was used to analyze the LUS score and lung consolidation to predict NRDS severity. They reported that the number of consolidation areas increased with NRDS severity (0 vs. 1.5 ± 0.8 vs. 4.1 ± 1.3, all $P < 0.05$). The LUS score for NRDS vs. non-NRDS showed 80.2% sensitivity and 100% specificity using a cut-off of 21.5 (Area under the ROC curve, AUC = 0.938; $P < 0.001$). They concluded that the LUS score and consolidation areas can discriminate NRDS from non-NRDS and the different grades of NRDS, and predict the application of mechanical ventilation.

Although this article has provided valuable information, but there are some substantial points that considering them can help the clarity of the method and an accurate interpretation of the study. First, it should be noted that diagnostic value of LUS are included the assessment of both validity (accuracy) and reliability (precision) of the test. The point is, due to the limitation of reported values for validity of a test (e.g. sensitivity and specificity are generally used for public health purposes and limited in clinical practice. PPV extremely depends on the prevalence of the outcome); other validity estimates such as likelihood ratios, should also be reported. Reported estimates as in this study can be acceptable; however, considering the rest of validity estimates, our final decision can be changed [2].

Second, without assessing reliability (precision), we cannot talk about diagnostic value of a test [2–7]. Third, receiver operative characteristics (ROC) curve is usually used to assess diagnostic accuracy (discrimination) of a diagnostic model. However, for clinical purposes, reporting diagnostic added value of LUS is crucially important. The reason is all validity estimates can be acceptable but diagnostic added value may be negligible [2–7].

Finally, for prediction, assessing internal and external validity is recommended. That is why we need two different set of cohort dataset [8,9]. It would be reasonable to assess interaction between predictors before developing a prediction score.

In the light of the mentioned points, I discussed common mistakes in assessing the diagnostic value of a test. I also emphasized on methodological issues on prediction studies. Any conclusion in diagnostic value and prediction needs to be supported by the methodological and statistical issues mentioned above. Otherwise, misinterpretation cannot be avoided.

Source(s) of support

None.

Declaration of Competing Interest

None.

References

[1] H. Pang, B. Zhang, J. Shi, J. Zang, L. Qiu, Diagnostic value of lung ultrasound in evaluating the severity of neonatal respiratory distress syndrome, Eur. J. Radiol. 116 (2019) 186–191.
[2] M. Szklo, F.J. Nieto, Epidemiology Beyond the Basics, 3rd. Manhattan, Jones and Bartlett Publisher, New York, 2014.
[3] S. Sabour, A common mistake in assessing the diagnostic value of a test: failure to account for statistical and methodologic issues, J. Nucl. Med. 58 (2017) 1182–1183.
[4] S. Sabour, F. Ghassemi, The validity and reliability of a signal impact assessment tool: statistical issue to avoid misinterpretation, Pharmacoepidemiol. Drug Saf. 25 (2016) 1215–1216.

DOI of original article: https://doi.org/10.1016/j.ejrad.2019.05.004

https://doi.org/10.1016/j.ejrad.2019.108663
Received 9 July 2019; Accepted 6 September 2019

[5] S. Sabour, Validity and reliability of the new Canadian Nutrition Screening Tool in the' real-world' hospital setting: methodological issues, Eur. J. Clin. Nutr. 69 (2015) 864.
[6] S. Sabour, Methodologic concerns in reliability of noncalcified coronary artery plaque burden quantification, AJR Am. J. Roentgenol. 203 (2014) W343.
[7] S. Sabour, E.V. Dastjerdi, Reliability of four different computerized cephalometric analysis programs: a methodological error, Eur. J. Orthod. 35 (2013) 848.
[8] S. Sabour, Prediction of 3-dimensional pharyngeal airway changes after orthognathic surgery: a methodological issue, Am. J. Orthod. Dentofacial Orthop. 147 (2015) 8.
[9] S. Sabour, Obesity predictors in people with chronic spinal cord injury: common mistake, J. Res. Med. Sci. 18 (2013) 1118.

Siamak Sabour[a,b,*]
[a] *Department of Clinical Epidemiology, School of Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran*
[b] *Safety Promotions and Injury Prevention Research Centre, Shahid Beheshti University of Medical Sciences, Tehran, Iran*
E-mail address: s.sabour@sbmu.ac.ir.

[*] Corresponding author at: Chamran Highway, Velenjak, Daneshjoo Blvd, Department of Clinical Epidemiology, School of Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran.