Evaluation of Early Warning Scores for Sepsis Prediction on the General Ward

Nirmala Shivakumar, BS\(^1\); Sean C. Yu, MS\(^1\); Andrew P. Michelson, MD\(^2\); Aditi Gupta, PhD\(^1\); Kevin Betthauser, PharmD\(^3\); Marin H. Kollef, MD\(^2\); Albert M. Lai, PhD\(^1\); Philip R.O. Payne, PhD\(^1\)

\(^1\)Institute for Informatics, Washington University School of Medicine, St. Louis, MO  
\(^2\)Division of Pulmonary and Critical Care, Washington University School of Medicine, St. Louis, MO  
\(^3\)Department of Pharmacy, Barnes-Jewish Hospital, St. Louis, MO

**Introduction**

Sepsis is life-threatening organ dysfunction due to a dysregulated host-response to infection\(^1\). Early identification and appropriate management of sepsis is vital to providing optimal care: delay in antimicrobial therapy can increase mortality by up to 7.6% per hour\(^2\). Early warning scores (EWS) for assessing clinical deterioration, commonly used in hospitals, are now being evaluated for sepsis prediction. Prior studies evaluating EWS to predict sepsis have been limited to the intensive care unit (ICU) and not formally evaluated on the general ward. In this research, we calculated EWS for patients on the general ward and assessed their predictive performance.

**Methods**

Clinical data was extracted from the electronic health record at Barnes-Jewish Hospital, a large tertiary-referral academic medical center in St. Louis, MO. All adult admissions to a non-ICU floor between 2012 and 2017 were eligible for inclusion. Admissions were excluded if they had no diagnosis code, no vital signs recorded, ICU transfer during or prior to prediction period, or if the admitting service was psychiatry or obstetrics/gynecology. Sepsis cases were identified using the Sepsis-2 criteria with suspicion for infection\(^3\). Sepsis onset was defined as the first time the criteria were met within a 24-hour window. For the non-sepsis population, the index time was defined as the midpoint between admission and discharge. Each EWS was calculated 6-hours before the index time using clinical data from the preceding 24 hours. The EWS evaluated included the National Early Warning Score (NEWS), National Early Warning Score 2 (NEWS 2), Modified Early Warning Score (MEWS), Systemic Inflammatory Response Syndrome (SIRS), Quick Sequential Organ Failure Assessment (qSOFA), and Sequential Organ Failure Assessment (SOFA). Missing data were imputed using normal physiologic values. All analyses were completed using Python and packages (i.e., NumPy, pandas, and scikit-learn).

**Results**

Of the 61,364 admissions, 20,868 met inclusion criteria, of which 471 (2.26%) met sepsis criteria. Area under the precision recall curve (AUPRC) ranged from 0.028 to 0.066 and area under the receiver operating characteristic (AUROC) ranged from 0.521 to 0.701 for all EWS (Figure 1). The median number of missing elements for EWS calculation were: 1 out of 7 for NEWS; 3 out of 9 for MEWS; 0 out of 4 for SIRS; 1 out of 3 for qSOFA; and, 3 out of 7 for SOFA.

**Conclusion**

The low AUPRC and AUROC suggest EWS measures are not reliable indicators for sepsis-specific prediction on the general ward. Missing data elements may have played a role in the unreliability of the EWS as predictors of sepsis, and showcase them as impractical measures for many hospital floors. Future analysis includes evaluation of temporal patterns, missing data subgroups and clinical outcomes, as well as comparisons of the EWS to machine-learning models for sepsis prediction.

**References**