Supplementary Online Content

Wong A, Otles E, Donnelly JP, et al. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Intern Med.* Published online June 21, 2021. doi:10.1001/jamainternmed.2021.2626

This supplementary material has been provided by the authors to give readers additional information about their work.

eMethods

The Epic Sepsis Model

The Epic Sepsis Model (ESM) is a penalized logistic regression model developed from a pooled sample of 405,000 patient encounters across three health care organizations between 2013 and 2015. Data was collected from the electronic health record in 30 minute observation intervals, up to 24 hours prior to the time of clinical intervention, defined as initiation of antibiotics, documentation of sepsis or suspicion of sepsis, usage of a sepsis-related order set, or an order for a lactate lab. Data elements included vital signs, medication orders, lab values, comorbidities, and demographic information. For model development, sepsis was defined as any encounter associated with an International Classification of Diseases (ICD-9) code indicating diagnosis of sepsis. Time of sepsis onset was defined as 6 hours prior to clinical intervention, with any time point falling prior to 6 hours before the time of clinical intervention labeled as negative for sepsis. Site-specific models were separately trained at each of the three institutions and the model coefficients were averaged to create a final 80-variable model. Model performance for the final model was separately assessed at each site, and the area under the receiver operating characteristic curve (AUC) ranged between 0.76 to 0.83.

Additional Details on the Definition of Sepsis and Timing of Onset

Patients who only had an International Classification of Diseases-10 diagnosis of sepsis but did not meet other objective criteria for sepsis were considered not to have sepsis. Among patients who met the 2 Systemic Inflammatory Response Syndrome and 1 Centers for Medicare and Medicaid Services organ dysfunction criteria within 6 hours, the later time was used to define the time of sepsis onset. Among patients only meeting the Centers for Disease Control clinical surveillance definition, sepsis onset was defined by the first time at which the definition was met.

Rationale for Use of the Hospitalization-Level AUC

The rationale for evaluating the AUC at the hospitalization level is that if a hypothetical alert were to be linked to a score threshold, whether the alert ever fired for any given patient would depend on whether this threshold was ever exceeded during the hospitalization. If a patient crossed a given alerting threshold even once prior to the outcome, this would bring the patient to the clinician’s attention if linked to an alert.

Calculation of Time Horizon-Based AUC

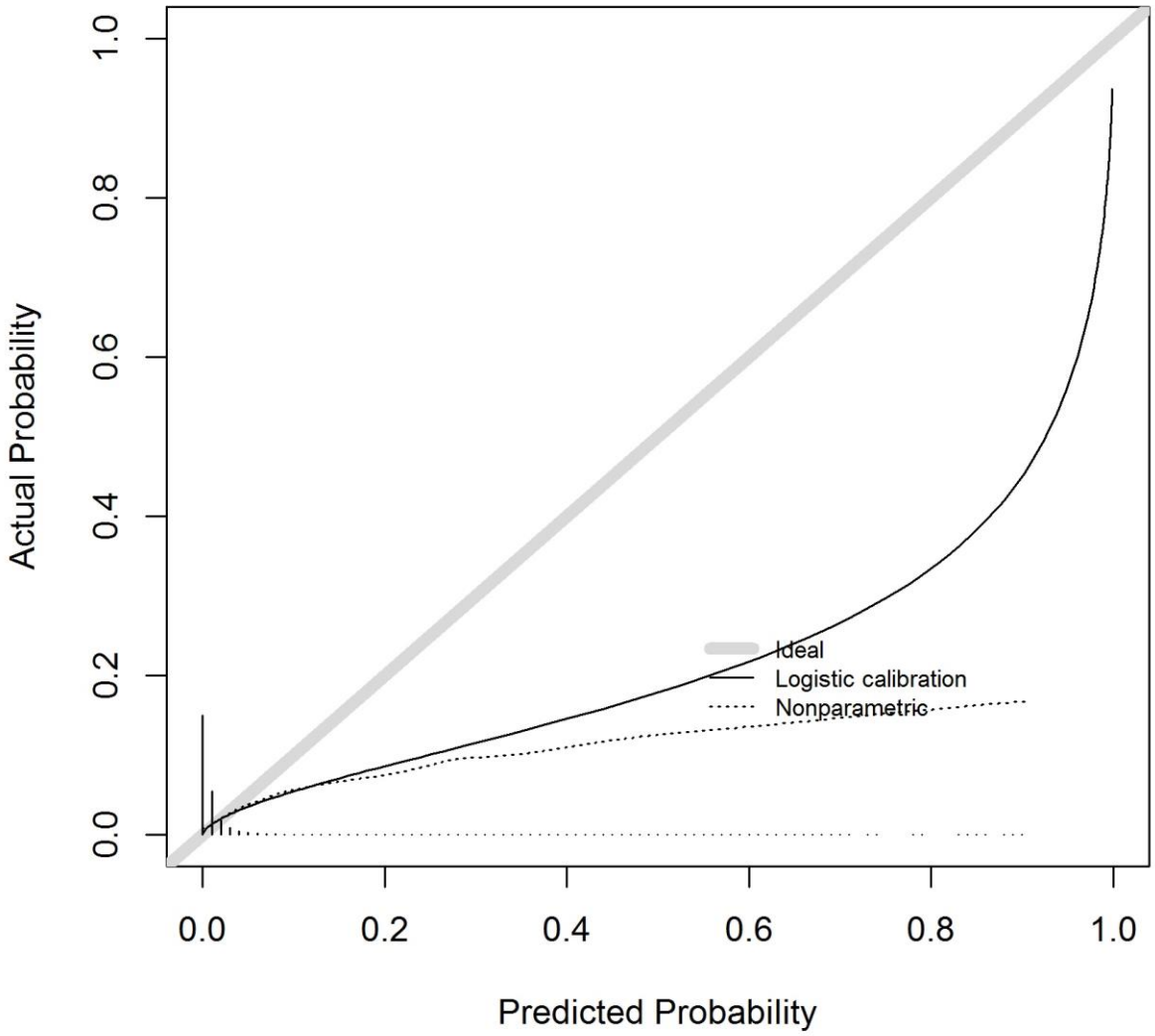
We also used a time horizon-based approach to calculate the AUC, which treats each prediction as completely independent and considers it as accurate if sepsis occurs within a given time horizon. This method has several limitations. First, it gives excess weight to patients with longer hospitalizations. Second, the scores for an individual are not actually independent, and thus even a small number of bad predictions (i.e., high scores that result in alerts in a non-septic patient) can cause alert fatigue despite having a minimal impact on the time horizon-based

AUC. Third, it ignores the fact that repeated positive predictions are not clinically relevant because repeat alerts are commonly “muted” after an initial alert. Although the time horizon-based approach has several limitations, it is commonly reported in the literature. Thus, we calculated AUCs with time horizons of 4, 8, 12, and 24 hours.

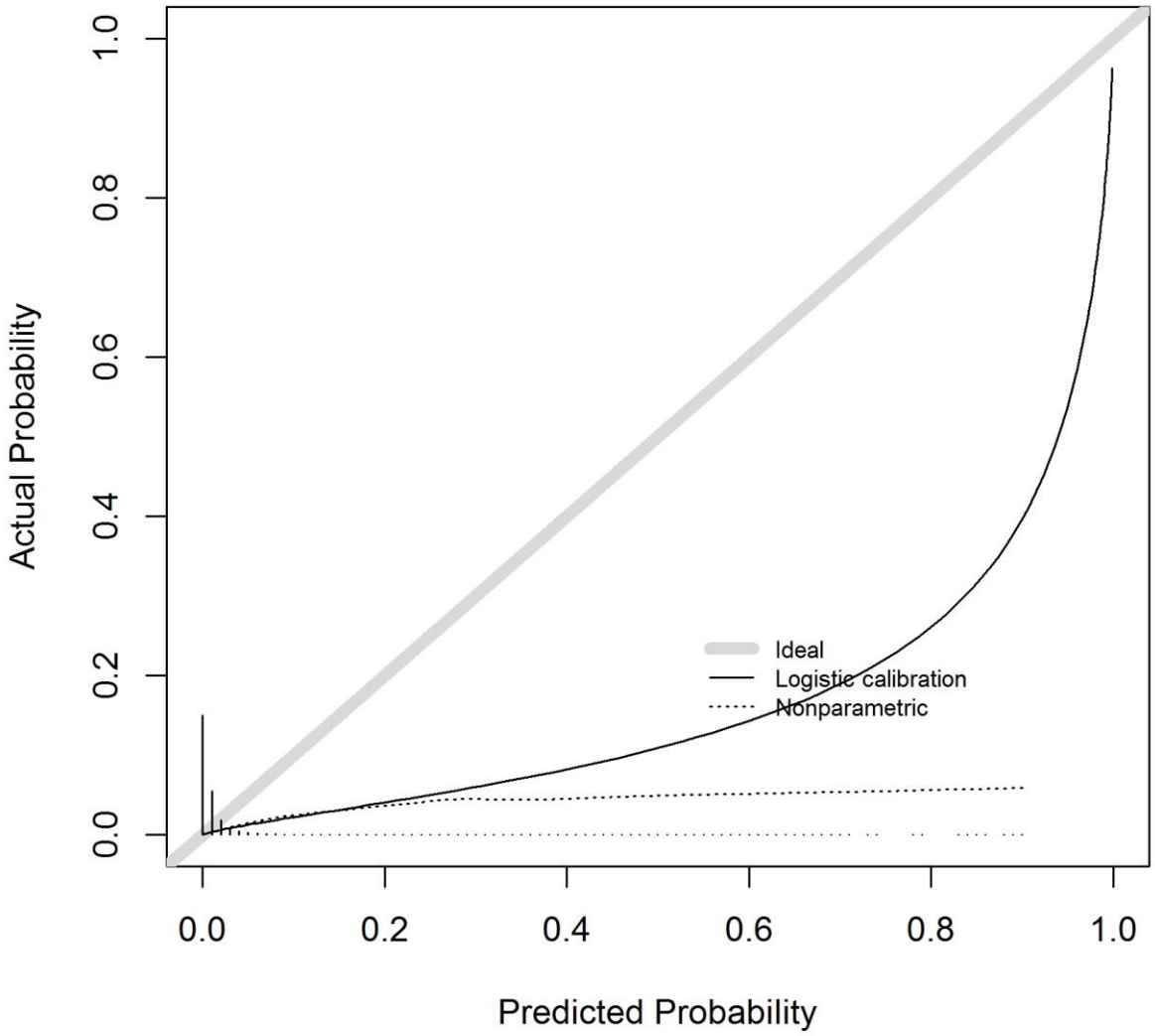
Rationale for Sensitivity Analysis

Prior evaluations of the ESM either did not exclude scores from after sepsis onset[1] or used clinical actions to define sepsis onset (unpublished Epic evaluation), which can bias the results in favor of the ESM. To evaluate the impact of this evaluation decision, we conducted a sensitivity analysis in which the model scores from up to 3 hours *after* sepsis onset were included in the evaluation.

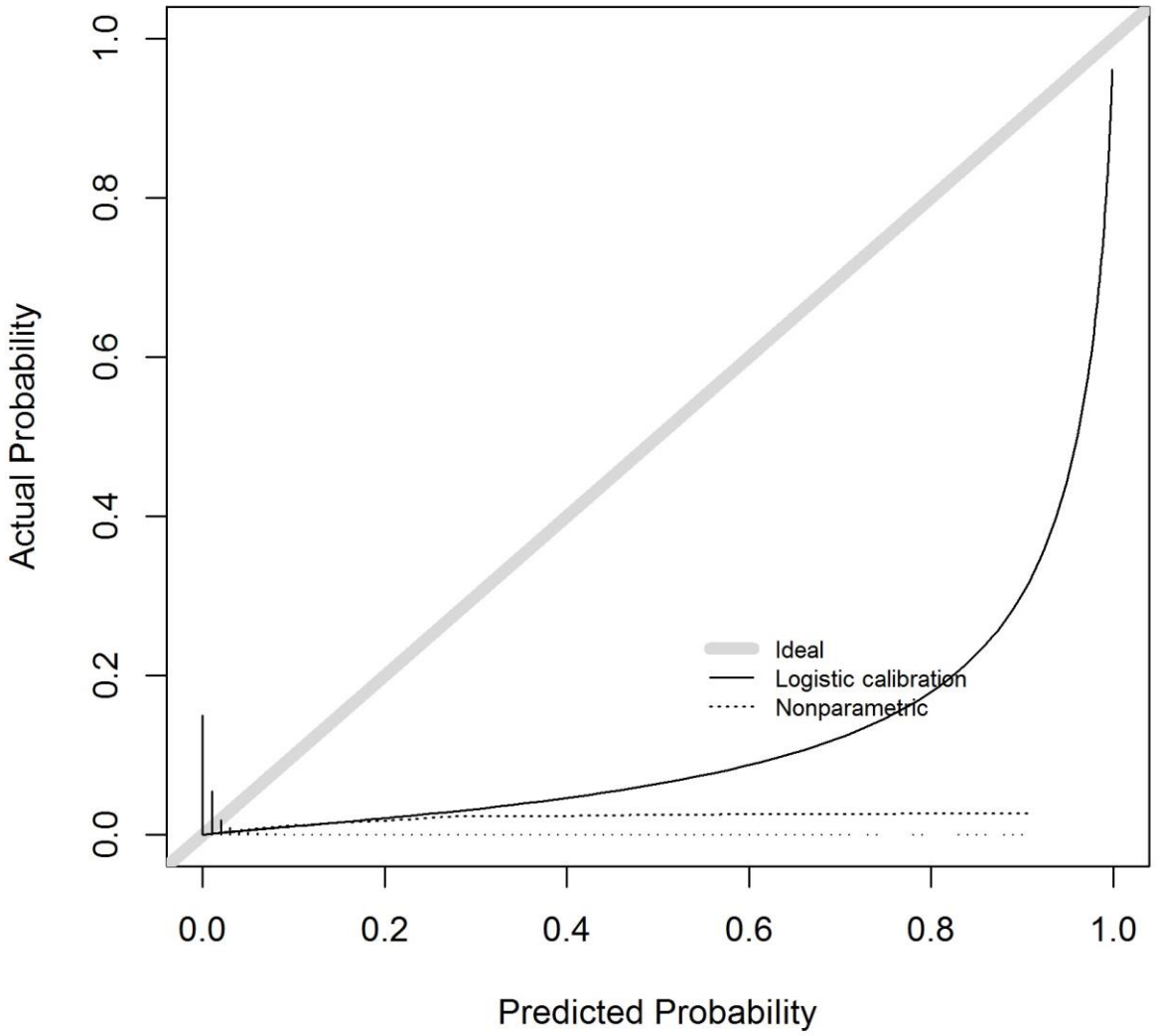
eFigure 1. Calibration Plot Comparing Continuous Estimates of Predicted ESM, Rescaled From 0-1, and the Observed Risk of Sepsis During the *Hospitalization*. Non-parametric refers to a lowess curve, while logistic calibration refers to a logistic regression model. A histogram of model predictions is placed above the x-axis.



eFigure 2. Calibration plot comparing continuous estimates of predicted ESM, rescaled from 0-1, and the observed risk of sepsis during the *next 24 hours*. Non-parametric refers to a lowess curve, while logistic calibration refers to a logistic regression model. A histogram of model predictions is placed above the x-axis.



eFigure 3. Calibration plot comparing continuous estimates of predicted ESM, rescaled from 0-1, and the observed risk of sepsis during the *next 6 hours*. Non-parametric refers to a lowess curve, while logistic calibration refers to a logistic regression model. A histogram of model predictions is placed above the x-axis.



eReference.

1. Bennett T, Russell S, King J, et al. Accuracy of the epic sepsis prediction model in a regional health system. *arXiv [statAP]*. Published online February 19, 2019. http://arxiv.org/abs/1902.07276