**Method**

***Sample Characteristics*.** The study also included measures to describe characteristics of the current sample that do not vary hour-by-hour.

This includes self-report measures of current *depressive symptoms* (total score of the Patient Health Questionnaire-9; Spritzer, Kroenke, & Williams, 1999) and past year *problematic alcohol use* (total score of 8 or greater on the Alcohol Use Identification Test; Saunders, Aasland, Babor, de la Fuente, & Grant, 1993). History of suicide attempts was assessed with the following question from the Military Suicide Research Consortium’s Common Data Elements (Littlefield et al., 2021): Including the suicide attempt that led to your current hospitalization, “How many times in your lifetime have you made an attempt to kill yourself during which you had at least some intent to die?” Thus *prior history of a suicide attempt* was coded (1) if the participant indicated that they made 2 or more attempts in their lifetime and was coded (0) if they indicated that this was their only attempt. Next, past year treatment history (prior to the recent hospitalization) was assessed by three questions from the WHO-Composite International Diagnostic Interview CIDI 3.0 (Kessler and Üstün, 2004) which assessed presence of past-year 1) overnight *psychiatric hospitalization* (“how many different times were you hospitalized overnight for your emotions, nerves, or mental health?”); 2) *psychotropic medication use* for your “emotions, nerves, or mental health from any type of professional”; and 3) having *outpatient psychological counseling* (“not counting the times you were an overnight patient in the hospital, did you have one or more sessions of psychology counseling or therapy for emotional problems that lasted 30 minutes of longer with any type of professional?”). Previous treatment variables were each coded as present (1) or absent (0).

Finally, participants reported the full date and time of their index suicide attempt that led to the current hospitalization and the date and time of the TLFB-SA was recorded by research staff. This information was used to determine number of days from the suicide attempt to study interview. Time of day for the suicide attempt was assessed by four, six-hour blocks: morning [6:00 AM-11:59 AM], afternoon [12:00 PM-5:59 PM], evening [6:00 PM to 11:59 PM], and overnight [12:00 AM-5:59 AM]). To assess suicide attempt method, participants were asked, “What method(s) did you use?” (Kessler & Üstün, 2004).

**Discussion**

We would like to highlight that AUC is threshold (i.e., the probability cut-value to determine when the model predicts a positive, here a case period) independent because it shows the sensitivity/specificity tradeoff at different thresholds (i.e., it summarizes predictive power for all possible probability thresholds; see Agresti, 2013, p. 224, section 6.3.4). Conversely, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) are threshold dependent (i.e., one can obtain different values for these metrics by changing the threshold, as is demonstrated on a ROC curve for sensitivity and 1-specificity; see Agresti, 2013, p. 225). We also note that the outcome, case (suicide attempt period) vs. control (no suicide attempt period), is a fixed, rather than a random, variable (i.e., by definition, 50% of the periods are case and 50% are control). Thus, metrics such as PPV and NPV, which are dependent on the base rate of the outcome (per Bayes theorem; see Efron, 2013), are impacted by these fixed periods. That is, despite the larger literature typically reporting very low PPV for outcomes of suicide attempt (given this outcome typically has a very low base rate when assessed as a random variable), the current data indicates a relatively high PPV (i.e., 0.74). Conversely, sensitivity and specificity (which are inherently part of the AUC analyses) are not driven by the base rate of the outcome (since the status of the outcome is given when calculating sensitivity and specificity; see Agresti, 2013, p. 39, section 2.1.3). Thus, the study design should be considered when interpreting PPV and NPV. These values may be misleading given the fixed outcomes in this design and are included in the current manuscript per a Reviewer’s request.

**Supplemental References**

Agresti, A. (2013). *Categorical Data Analysis.* 3rd Edition, John Wiley & Sons Inc., Hoboken.

Efron, B. (2013). Bayes' theorem in the 21st century. *Science*, *340*(6137), 1177-1178.

Kessler, R. C., & Ustun, T. B. (2004). The World Mental Health (WMH) Survey Initiative

Version of the World Health Organization (WHO) Composite International Diagnostic

Interview (CIDI). *International Journal of Methods in Psychiatric Research, 13*(2), 93-121. doi: 10.1002/mpr.168

Littlefield, A.K., Cooke, J.T., Bagge, C., Glenn, C., Kleiman, E.M., Jacobucci, R., … Steinley, D. (2021).

Machine learning to classify suicidal thoughts and behaviors: Implementation within the common data

elements used by the military suicide research consortium. *Clinical Psychological Science, 9*(3),467-

481. doi:10.1177/2167702620961067

Saunders, J.B., Aasland, O.G., Babor, T.F., de la Fuente, J.R., & Grant, M. (1993). Development of the Alcohol

Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons

with Harmful Alcohol Consumption--II. *Addiction, 88*(6),791-804. doi:10.1111/ j.13600443.1993. tb02093.x

Spitzer, R.L., Kroenke, K., Williams, J.B. (1999). Validation and utility of a self-report version of PRIME-MD: The PHQ primary care study. Primary Care Evaluation of Mental Disorders. Patient Health Questionnaire. *Journal of the American Medical Association, 282*(18),1737-1744. doi:10.1001/jama.282.18.1737