Chronic pain following traumatic stress is common in the US. Although most individuals recover following traumatic stress exposure, a substantial proportion develop persistent severe pain. This project aimed to develop a simple-to-use and highly accurate tool to predict the development of severe pain after traumatic stress. Modern methods often use complex analytical techniques such as machine learning to make highly accurate predictions using many features. These methods are difficult to implement quickly with limited computational resources and specialized training, hindering their use in fast-paced settings like the emergency department where people often report following traumatic experiences. Simpler models, such as regression, while unable to consider complex relationships between predictor and outcome variables, are easier to implement quickly.

Methods
Study Design and Setting—Data used in the current study (see Table 1 for demographics) was a part of AURORA study (Figure 1), a national multi-site prospective study based in 30 US emergency departments (EDs). The aim of the AURORA study is to gain insight into adverse posttraumatic neuro-psychiatric sequelae among trauma survivors, including pain.

Participants—AURORA participants were included in the study if they reported to the ED within 72 hours following a motor vehicle collision (MVC)-related traumatic event, were not admitted to the hospital, and completed both 2-week and 3-month assessments.

Measurements—Predictor measures included 265 variables spanning demographic, psychological and personality traits, past experiences/stressors, and physical health categories. The outcome measure was severe pain, defined as Pain Numeric Rating Scale (NRS) scoring ≥7. Pain NRS was scored on a 0-10 scale where 0 indicated no “pain or tenderness” and 10 represented “severe pain or tenderness.”

Data Analysis—Ten lasso logistic regressions in randomly selected (bootstrapped) cohort subsamples were performed to determine the top 20 predictors (Figure 2) based on regression coefficient. Then, each predictor was converted into binary variables based on dichotomizing each level of response options (Table 2). The final lasso logistic regression model was developed based on the selected number of binarized variables. Model performance (Table 3) was assessed considering both discrimination (i.e., area under the receiver operating characteristic curve [AUC]) and accuracy of predicted risk probabilities (i.e., Brier).