Development of Machine Learning Algorithms for 90-Day and 1-Year Mortality Prediction in the Elderly with Femoral Neck Fractures based on the HEALTH and FAITH Trials

*Hidde Dijkstra, BS; **Jacobien Oosterhoff MD**; Anouk Van De Kuit, BS; Frank Ijpma, MD, PhD; Joseph Hasbrouck Schwab, MD, FAAOS; Rudolf W. Poolman, MD, PhD; Sheila Sprague, PhD; Sofia Bzovsky, MSc; Mohit Bhandari, MD, FRCSC; Marc F. Swiontkowski, MD; Emil H. Schemitsch, MD, FAAOS; Job N. Doornberg, MD, PhD; Laurent Alexander Marijn Hendrickx, MD*
University Medical Centre Groningen, Groningen, NETHERLANDS

Purpose: Preoperative estimation of mortality is crucial for treatment decision-making for the elderly with a femoral neck fracture (FNF). Machine Learning (ML) algorithms might improve mortality estimation and are ideally developed with high-quality prospective data. The purpose of this study was to develop ML algorithms for predicting 90-day and 1-year mortality rates using the HEALTH (Hip fracture Evaluation with Alternatives of Total hip arthroplasty versus Hemiarthroplasty) and FAITH (Fixation using Alternative Implants for the Treatment of Hip fractures) data sets.

Methods: Patients from the HEALTH and FAITH trials were randomly divided into a training set (80%) and hold-out set (20%). Subsequently, feature selection was carried out by random forest algorithms. Based on the selected features, 6 algorithms were used on the training set using 10-fold cross-validation repeated 3 times and the 5 best performing were subsequently tested in the hold-out set. The best performing algorithms for 90-day and 1-year mortality prediction, evaluated across discrimination (c-statistic), calibration (slope and intercept) and the Brier score, were incorporated into an open-access web application.

Results: This study included 2388 patients with a 90-day mortality rate of 3.0% (n = 71) and 1-year mortality rate of 6.4% (n = 153). Nine variables were identified as best predictors for 90-day mortality with prefracture functional status (the use of ambulatory assistive device), American Society of Anesthesiologists (ASA) classification, sex ,and osteoarthritis being the 4 most predictive. For predicting 1-year mortality, we identified 15 predictive variables with ASA classification, prefracture functional status, ethnicity, and age being the 4 most predictive. The best-performing algorithm was the penalized logistic regression algorithm for 90-day mortality ([metrics noted as training set/hold-out set] c-statistic: 0.78/0.80, calibration slope: 0.93/0.95, calibration intercept: 0.0/–0.06, and Brier score: 0.039/0.039) and for 1-year mortality (c-statistic: 0.78/0.76, calibration slope: 0.95/0.86, calibration intercept 0.00/–0.20, and Brier score 0.73/0.74).

Conclusion: We developed robust ML-derived prediction models for mortality estimation for elderly patients with FNFs based on patient and fracture characteristics. This ML probability calculator may facilitate data-driven shared decision-making: empowering patients and their families to make a personal decision for a surgical strategy that best fits their individual values and needs. The open-access web application can be found at: https://traumaplatform-ai-prediction-tools.shinyapps.io/90daymortalityhip/ and https://traumaplatform-ai-prediction-tools.shinyapps.io/1yearmortality/