

## Development and Internal Validation of a Clinical Prediction Model Using Machine Learning Algorithms for 90-Day and 2-Year Mortality in Femoral Neck Fracture Patients Aged 65 Years or Above

Jacobien Oosterhoff, MD; Berit Marte Corlijn Savelberg, BS; Aditya Vishwas Karhade, MD; Benjamin Gravesteijn, MA; Job N. Doornberg, MD; Joseph Hasbrouck Schwab, MD; Marilyn Heng, MD

Massachusetts General Hospital, Boston, Massachusetts, UNITED STATES

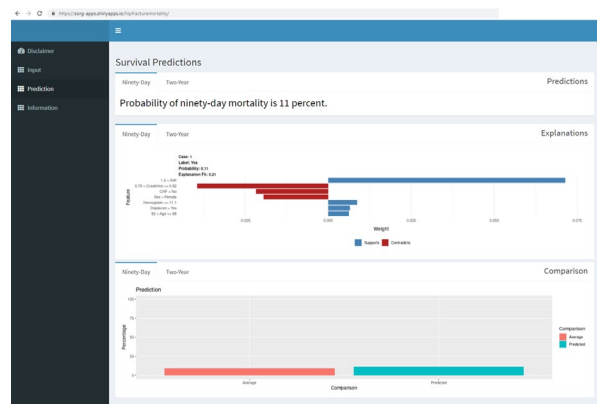
**Purpose:** Predicting mortality in elderly femoral neck fracture patients can improve the challenging surgical decision-making. A preoperative clinical prediction model may aid surgeons in this challenging treatment process, and optimize care for elderly femoral neck fracture patients. Therefore we questioned, can we develop and internally validate a clinical prediction model for 90-day and 2-year mortality in femoral neck fracture patients aged 65 years or above?

**Methods:** A retrospective cohort study at 2 Level I trauma centers and 3 community hospitals was conducted to identify patients undergoing surgical fixation for a femoral neck fracture. Variable selection was performed to identify and select those preoperative variables contributing most to predicting mortality, conducted by entering all relevant explanatory variables into random forest algorithms with recursive selection. Five different machine learning (ML) algorithms were developed and internally validated and assessed by discrimination, calibration, Brier score, and decision curve analysis.

**Results:** In total, 2478 patients were included with 90-day and 2-year mortality rates of 9.1% and 23.5%, respectively. The models included patient characteristics, comorbidities, and laboratory values. The stochastic gradient boosting algorithm had the best performance for 90-day mortality prediction, with good discrimination (c-statistic = 0.74), calibration (intercept = -0.05, slope = 1.11), and Brier score (0.078). The elastic-net penalized logistic regression algorithm had the best performance for 2-year mortality prediction, with good discrimination (c-statistic = 0.70), calibration (intercept = -0.03, slope = 0.89), and Brier score (0.16). The models were incorporated into an open-access web-based application, including individual patient explanations for understanding how the model reached a certain prediction, and overcome the “black box”:

<https://sorg-apps.shinyapps.io/hip-fracturemortality/>.

**Conclusion:** The clinical prediction models show promise in estimating mortality prediction in elderly femoral neck fracture patients. External and prospective validation of the models may improve surgeon ability when faced with challenging surgical decision-making.



The FDA has stated that it is the responsibility of the physician to determine the FDA clearance status of each drug or medical device they wish to use in clinical practice.