20294

The Preoperative Machine Learning Algorithm For Extremity Metastatic Disease Can Predict 90-Day and 1-Year Survival:

An External Validation Study Of 264 Patients

Mary K. Skalitzky BA[1], Trevor R. Gulbrandsen MD[1], Olivier Q. Groot BS[2], Aditya V. Karhade MD MBA[2], JJ Verlaan MD PhD, Joseph H. Schwab MD FAAOS[2], Benjamin J. Miller MD MS FAAOS[1]

[1]University of Iowa Hospitals and Clinics Dept. of Orthopedics & Rehabilitation; [2]Massachusetts General Hospital Dept. of Orthopedic Surgery; [3]University Medical Center Utrecht Dept of Orthopedics

Objective

Does the Skeletal Oncology Research Group (SORG) algorithm accurately predict 90-day and 1-year survival in an independent sample from an external institution surgically treated for metastatic long-bone disease in the extremities?

Introduction

The prediction and estimation of survival is valuable in determining the need for operative treatment of metastatic long-bone disease in the extremities.

A machine-learning algorithm that formulates survival predictions for patients with associated metastatic long bone disease has been previously developed and internally validated, with good performance.

However, external validation consisting of a separate population has not been previously performed.

Materials and Methods

Retrospective cohort study on patients who underwent operative management for long-bone metastasis

- **One academic institution: 2003-2019**

Variables used in the SORG algorithm:

- **Age**
- **Sex**
- **Primary tumor type**
- **Visceral metastases**
- **Brain metastases**
- **Previous systemic therapy**
- **10 preoperative laboratory values**

A previously validated stochastic gradient boosting model was used (prediction algorithm)

- **Model performance calculated by discrimination, calibration, and overall performance.**

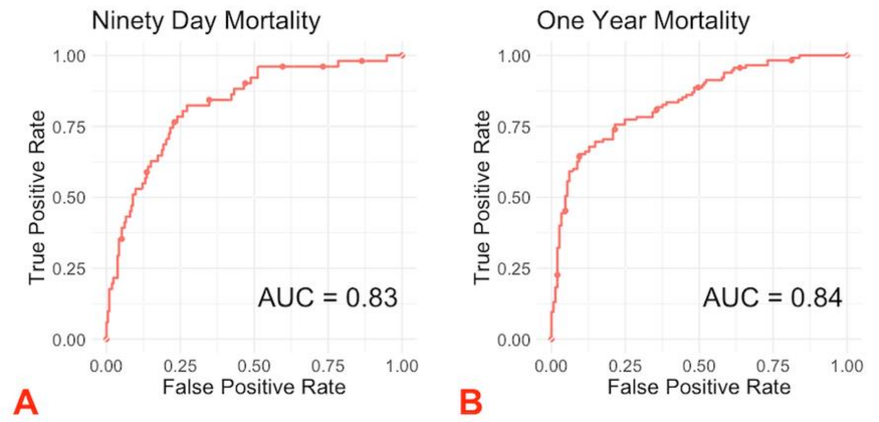
Results

A total of 264 patients

- **Median age of 64 years (IQR: 54-71)**
- **52% female**
- **All had > 1 year follow up**

The most common primary tumors included:

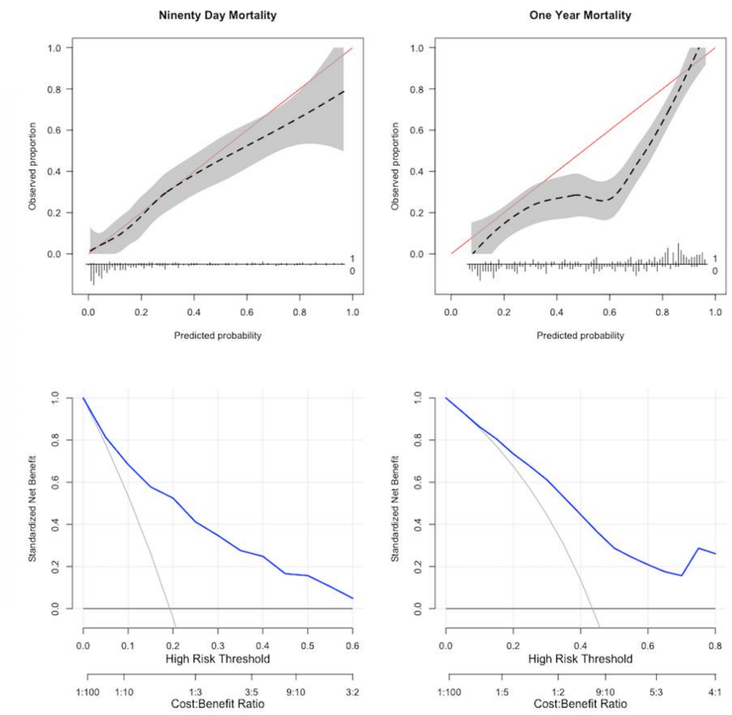
- **Renal cell (18%)**
- **Lung (16%)**
- **Multiple Myeloma (14%)**



The mortality (by any cause) was 19% (51/264) within 90-days and 42% (110/264) within 1-year.

Metric	90-day mortality	One-year mortality
	Discrimination	
AUC	0.83 (0.76, 0.88)	0.84 (0.79, 0.88)
F1-score*	0.56 (0.44, 0.67)	0.72 (0.63, 0.81)
	Calibration	
Intercept	-0.21 (-0.58, 0.17)	-0.73 (-1.02, -0.44)
Slope	0.84 (0.59, 1.09)	1.08 (0.81, 1.35)
	Overall performance	
Brier score	0.12 (0.10, 0.15)	0.18 (0.16, 0.21)
Null-model Brier score	0.16	0.25

AUC: area under the receiver operating curve. *probability threshold equal to the Youden index (90-day mortality threshold = 0.19, one-year mortality threshold = 0.76)



The validation cohort differed significantly from the developmental cohort on primary tumor histology, previous systemic therapy, and 1-year survival.

Despite these differences, the SORG ML algorithms generalized well to the validation cohort on discrimination (c-statistic 0.76−0.88 for 90-day mortality and 0.79−0.88 for 1-year mortality), calibration, Brier score, and decision curve analysis.

Discussion and Conclusions

The previously developed machine learning algorithms demonstrated good performance in this study, providing external validation.

These models are incorporated into an accessible application that may be freely utilized by clinicians in predicting survival for individual patients and assist in informative decision-making discussion with the patients prior to operative management of long bone metastatic lesions.

- **Application found at: https://sorg-apps.shinyapps.io/extremitymetssurvival/**