Supplemental Online Content

Chalkidis G, McPherson JP, Beck A, et al. External validation of a machine learning model to predict 6-month mortality for patients with advanced solid tumors. *JAMA Netw Open.* 2023;6(8):e2327193. doi:10.1001/jamanetworkopen.2023.27193

eMethods.

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods

To perform the external validation of the machine learning model to predict 6-month survival, we used all the methods implemented to develop and internally validate the original model (e.g., to identify the study cohort, extract records from the University of Utah enterprise data warehouse (EDW), extract treatment decision points (TDPs), etc.), but applied the model to an updated dataset of patient records. The model was originally developed using data for patients with advanced cancer that included TDPs between June 1, 2014 and June 1, 2020. During the external validation phase, we used data for patients who had TDPs between June 2, 2020 and April 12, 2022. This data included previously unseen patients who had an advanced cancer diagnosis after June 1, 2020 with subsequent lines of therapy (LoT), in addition to patients who were diagnosed with advanced cancer and had TDPs in the development phase but went on to receive new LoTs after November 30, 2020.

The data describing the study population demographics was directly extracted from electronic health record (EHR) data that is stored in the University of Utah EDW. Data regarding race, ethnicity, and sex were documented during routine clinical care. A race category of 'Other' was used to aggregate patients with 'Unknown' race and other subcategories accounting for <2% of the total sample population.

As described in our earlier publication[2], the following definitions and methods were used to perform the external validation. Advanced cancer was defined as malignant brain or nervous system cancer or any other solid tumor with metastases. The first available advanced cancer diagnosis in the EHR was defined as the index date. Patients 18 years or older on the index date for advanced cancer, with no history of hematologic malignancy

or bone marrow transplant who received at least one line of anticancer therapy, were eligible for this study. An eligible line of therapy (LoT) included chemotherapy, biologics, targeted therapy, immunotherapy, and hormonal therapy. Eligible TDPs (e.g., a visit within 30 days before a LoT start date) within the external validation dataset were identified using the same data processing techniques used during model development. The prediction outcome of the extreme gradient boosting machine learning (ML) algorithm was 6-month mortality after a TDP. For each TDP, we had a minimum follow-up of 6 months to assess a patient's status of either deceased or alive.

Development and internal validation of the ML-model are described in detail in our previous publication.[2] The ML-model uses 45 input features: current LoT, patient's age, body mass index, pain score, time since index date, time since past treatment, 12 laboratory results, 15 cancer types, 7 metastasis sites, and 5 three-month percent change observations (weight, albumin, alkaline phosphatase, hemoglobin, platelets). Since the extreme gradient boosting ML-model can handle missing data, no imputation was performed. Data after the TDP was not included in the ML-model, with the exception of data to establish the primary outcome.

Since each TDP was considered an independent observation, multiple observations per patients were used, which is a legitimate strategy in prognostic model validation as described in our previous publication.[2] The original ML-model was neither updated nor re-trained using external validation data before performance was evaluated on external validation data. No further model updating (e.g., recalibration) was performed. Additionally, to assess potential algorithm bias and data leakage, we evaluated model performance on subsets of the external validation dataset: a) excluding any patient who

had earlier TDPs in the development dataset, and b) using non-parametric bootstrap, randomly selecting only one LoT per patient.

The Institutional Review Board at the University of Utah classified this study as exempt, and no informed consent was required.