ABSTRACT

Colorectal cancer with liver metastasis (Colorectal Liver Metastases/CRLM) is the second leading cause of cancer-related death worldwide, with approximately 50% of patients developing liver metastases during the course of the disease. Accurate prediction of survival time is crucial for supporting more precise and personalized clinical decision-making. This study conducted a survival analysis of colorectal cancer patients with liver metastasis using the XGBoost-AFT model based on radiomic features extracted from CT imaging. The XGBoost-Accelerated Failure Time (AFT) model was integrated with Isotonic Regression calibration and compared against the standard AFT model based on the log-logistic distribution. The dataset was obtained from The Cancer Imaging Archive (TCIA) and included 197 patients with CRLM. Radiomic features were extracted using PyRadiomics and reduced through a two-stage feature selection process: ElasticNet-Cox followed by ensemble score-based selection (Pearson correlation, mutual information, F-score, and C-index). The XGBoost-AFT model was optimized using Bayesian hyperparameter tuning with 5-fold cross-validation. Evaluation on the test set showed that XGBoost-AFT outperformed the standard AFT model with a C-index of 0.83 vs. 0.79, IBS of 0.10 vs. 0.11, MAE Pseudo Observation of 20.98 vs. 22.88 months, and MAE event of 17.57 vs. 26.00 months. Kaplan-Meier curve analysis confirmed a highly significant risk stratification ability (log-rank $p = 3.33 \times 10^{-10}$). The XGBoost-AFT model successfully captured non-linear relationships among radiomic features and provided more accurate survival time estimates compared to conventional parametric models, demonstrating strong potential to support the implementation of precision oncology in the clinical management of CRLM patients.

Keywords: Colorectal Cancer, Liver Metastasis, Radiomics, Survival Analysis, Xgboost-AFT.