Supplementary

Table S1 Label encoding values for variables used in XGBoost models

Variables	Label encoding value				
	1	2	3	4	5
Sex	Male	Female	/	/	/
Race	White	Black	Other	/	/
Stage	In situ	Localized	Regional	Distant	/
Histologic type	Transitional cell carcinoma	Adenocarcinoma	Squamous cell carcinoma	Neuroendocrine carcinoma	Other epithelial tumors
Radiation therapy	Yes	No/Unknown	/	/	/
Chemotherapy	Yes	No/Unknown	/	/	/
Time from diagnosis to treatment	<1 month	1-3 months	Over 3 months	/	/
Tumor size	<30 mm	30-59 mm	60+ mm	/	/
Marital status	Married	Separated	Divorced	Widowed	Unmarried
Median household income	<$50,000	$50,000-100,000	$100,000+	/	/
Year of diagnosis	2000-2005	2006-2010	2011-2015	2016-2021	/

Table S2 Main parameters of the XGBoost model

Parameter	Value
Gama	1
Eta	0.1
Max_depth	6
Min_child_weight	10
Subsample	0.7
Colsample_bytree	0.8
nrounds	93 (6-month model)
	53 (12-month model)
	41 (36-month model)
	53 (60-month model)

Table S3 Performance of prognostic models built by machine learning algorithms on validation cohort (AUC value)

	6-month CSM	12-month CSM	36-month CSM	60-month CSM
XGBoost	0.799	0.756	0.746	0.745
Logistic regression	0.793	0.746	0.737	0.737
Random forest	0.779	0.748	0.731	0.740
Support vector machine	0.774	0.746	0.729	0.731

CSM: cancer specific mortality.

https://dx.doi.org/10.21037/tcr-24-2023

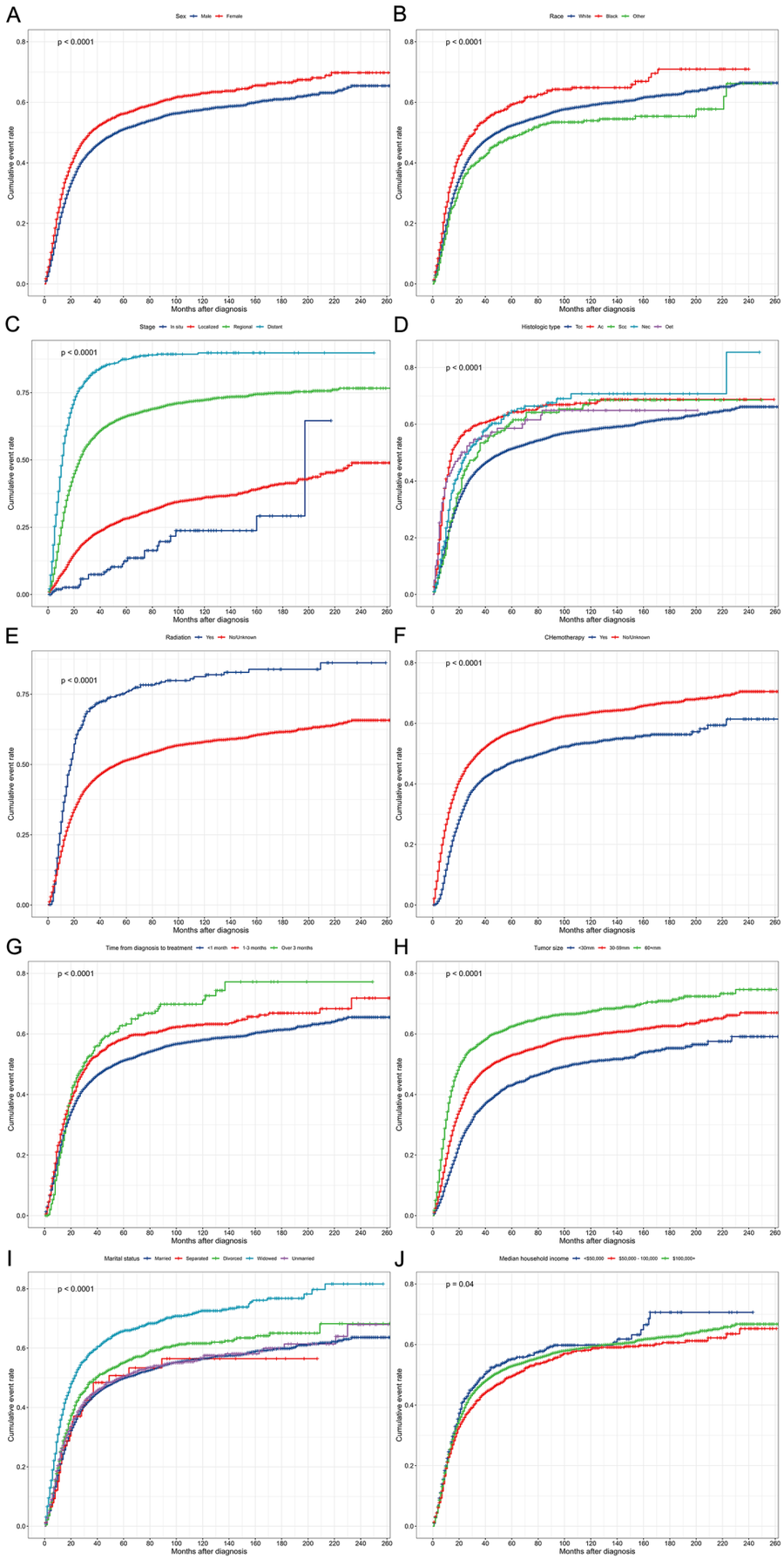

Figure S1 Cumulative mortality curves for bladder cancer-related deaths in patients with bladder cancer stratified by sex (A), race (B), tumor stage (C), histologic type (D), radiation therapy (E), chemotherapy (F), time from diagnosis to treatment (G), tumor size (H), marital status (I), and median household income (J). (Tcc: transitional cell carcinoma; Scc: squamous cell carcinoma; Nec: neuroendocrine carcinoma; Ac: adenocarcinoma; Oet: other epithelial tumors).

https://dx.doi.org/10.21037/tcr-24-2023

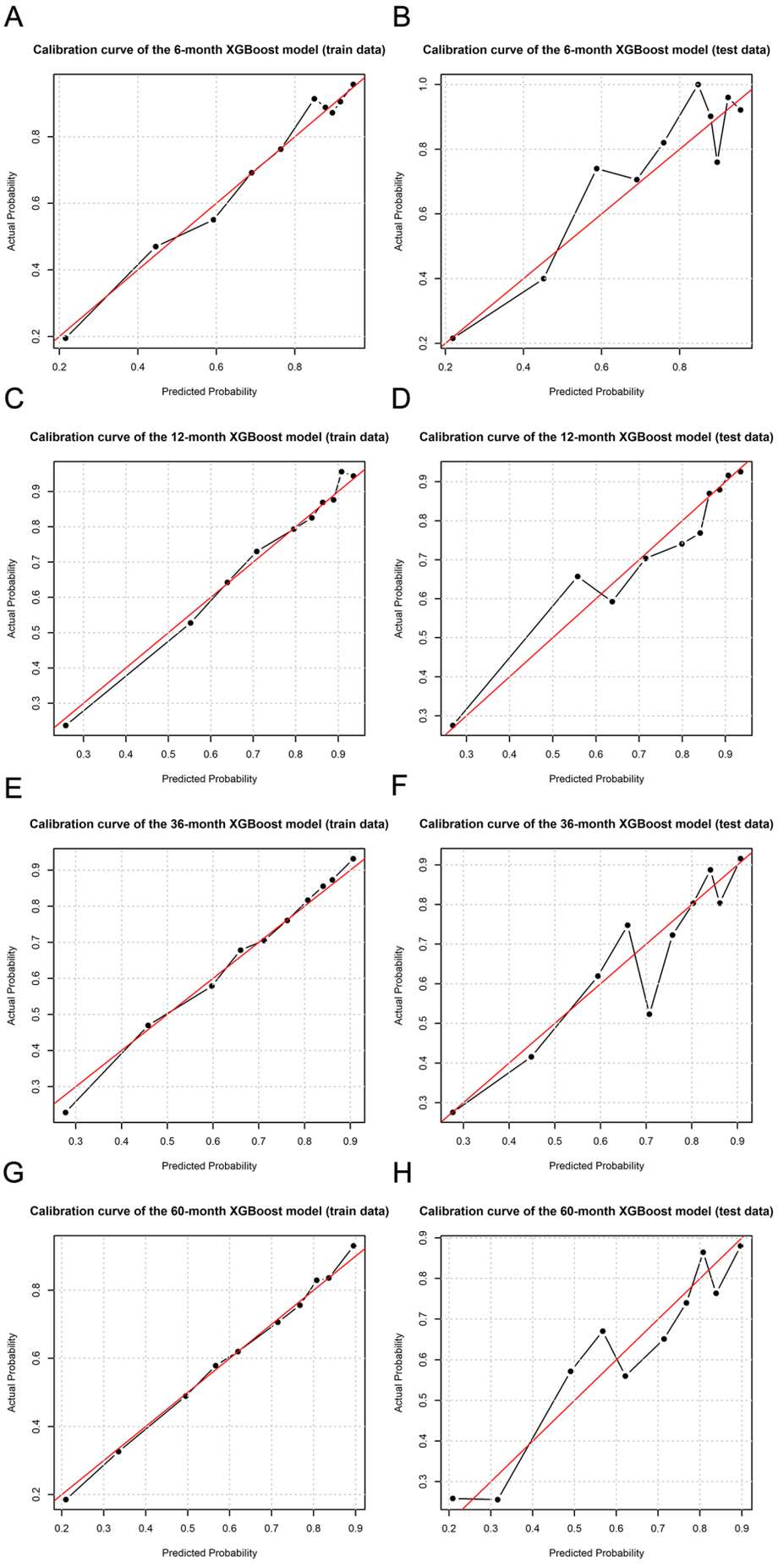

Figure S2 Calibration curves for XGBoost model evaluation: train data (A,C,E,G), test data (B,D,F,H). (XGBoost: extreme gradient boosting)

https://dx.doi.org/10.21037/tcr-24-2023

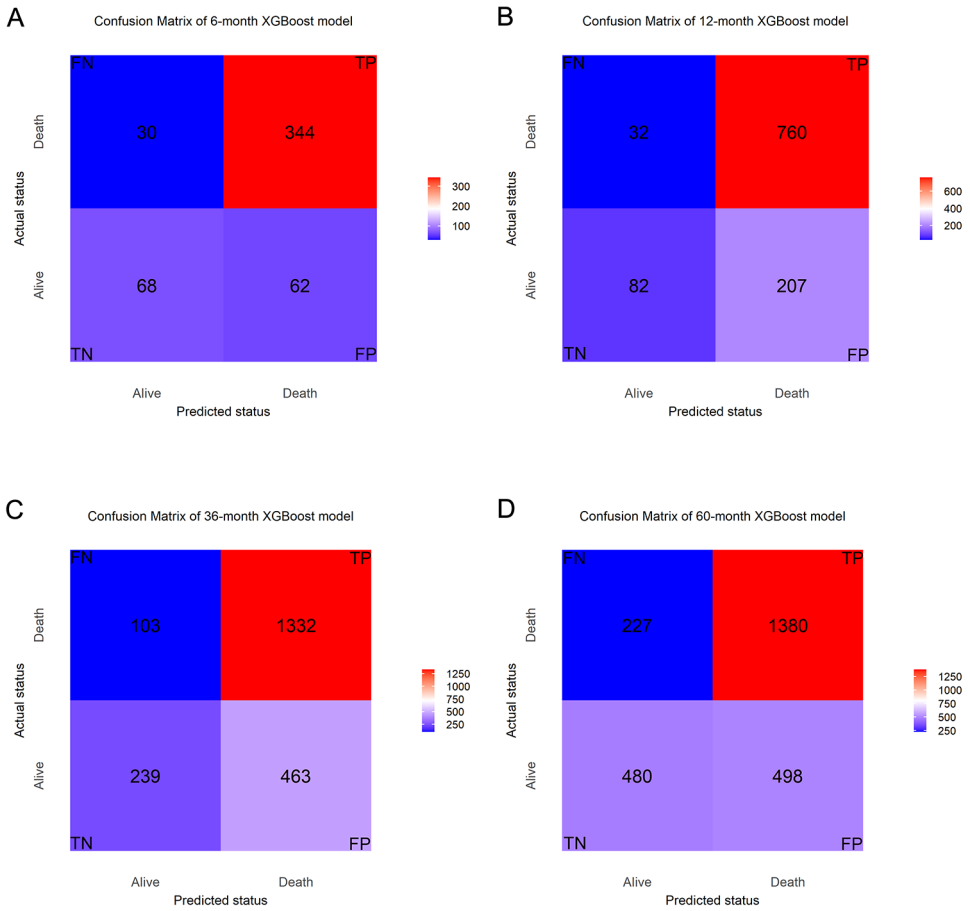

Figure S3 Confusion matrix of the XGBoost model. (XGBoost: extreme gradient boosting)

https://dx.doi.org/10.21037/tcr-24-2023