Supplemental Materials

Development and validation of a nomogram to predict the benefit of adjuvant radiotherapy for patients with resected gastric cancer

Supplementary Methods

Statistical analysis

MissForest is a random forest-based method that is able to efficiently handle missing data imputation among multivariate data by producing a single imputed dataset without setting aside test data or performing cross-validations.[1] To infer missing values regarding race (12 cases, 0.2%), tumor location (2,300 cases, 32.0%), tumor size (937 cases, 13.0%) and tumor differentiation (253 cases, 3.5%) in the SEER cohort, multiple imputations using missForest were performed with the following variables: year of diagnosis, patient age, gender, T stage, metastatic lymph node count (MLN), negative lymph node count (NLN), and receipt of ART.

We classified patients in the SEER cohort with the radiation code of "beam radiation" into the surgery+ART group (ART group) and those with the codes of "none" and "refused" into the non-ART group. Inverse probability propensity score weighting[2] was used to balance patient characteristics between the ART and non-ART groups among the training set and the SEER validation set. To calculate propensity scores, baseline covariates (patient age, year of diagnosis, race, tumor location, size, differentiation, T stage, MLN, and NLN) were applied to a logistic model for the receipt of ART. Based on the propensity score, each patient was weighted by the inverse probability of receiving ART, thus generating weighted synthetic samples in which observed baseline covariates were not confounded with ART assignment.[2]

OS was estimated using the Kaplan-Meier method. Multivariate Cox regression models with robust sandwich variance estimators were used to assess the relationships

between covariates and OS in the weighted samples.[3] Restricted cubic splines were used to examine functional forms of continuous variables in relation to survival.[4] Transformation using multivariable fractional polynomials was performed when the relationship was apparently non-linear;[5] the optimal transformation was obtained based on the Bayesian Information Criteria.[6] Multivariable fractional polynomial interactions were used to handle the interactions of ART with continuous variables.[7] For model construction, we began with the receipt of ART, other accounted variables (patient age, year of diagnosis, race, tumor location, size, differentiation, T stage, MLN, and NLN), and the first-order interaction terms between ART and other accounted variables. The final Cox models were obtained by using backward stepwise selection of the variables (keeping only those with $P<0.05$). The only treatment interaction term retained in the final model was with MLN, but this interaction was non-significant in the validation set ($P_{interaction}=0.67$). By contrast, the effect size of ART significantly varied by NLN in the SEER validation set ($P_{interaction}<0.01$). When MLN was replaced with the lymph node ratio[8] (LNR, defined as the ratio of MLN relative to the total examined nodes) in the model, a significant interaction between ART and LNR was detected in both the training set ($P_{interaction}<0.01$) and the SEER validation set ($P_{interaction}=0.01$), without diminishing the discriminatory abilities of the models. Therefore, we used LNR instead of MLN and NLN for model development.

A nomogram was developed to predict the 3- and 5-year OS probabilities given the LNR-based final model. Concordance indices (C-indices) were used to compare the discriminative abilities of the nomogram and the 8th AJCC staging system (i.e. the model including the 8th AJCC T and N classifications).[9] The value of the C-index ranges from 0.5

to 1.0, with 0.5 indicating a random chance and 1.0 indicating a perfect ability to correctly discriminate the outcome with the nomogram. Calibrations were performed by reviewing the plots of nomogram-predicted survival probabilities with the Kaplan-Meier-estimated probabilities.[10] Bootstraps with 1,000 resamples were used to quantify model overfit and calculate Kaplan-Meier estimates. External validation of the nomogram was carried out by discrimination and calibration using the SEER validation set .

Statistical significance was set as $P<0.05$ in a two-tailed test. The statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC, USA), SPSS v.19.0 (SPSS, Chicago, IL, USA), and R v.3.2.3 (http://www.r-project.org).

References

1. Stekhoven DJ, Buhlmann P. MissForest--non-parametric missing value imputation for mixed-type data. Bioinformatics. 2012; 28: 112-8.

2. Austin PC. The performance of different propensity score methods for estimating marginal hazard ratios. Stat Med. 2013; 32: 2837-49.

3. Lin DY, Wei LJ. The robust inference for the Cox proportional hazards model. J Am Stat Assoc. 1989; 84: 1074-8.

4. Royston P, Sauerbrei W. Multivariable modeling with cubic regression splines: a principled approach. Stata J. 2007; 7: 45-70.

5. Sauerbrei W, Royston P. Corrigendum: building multivariable prognostic and diagnostic models: transformation of the predictors by using fractional polynomials. J R Stat Soc Ser A Stat Soc. 2002; 165: 399-400.

6. Volinsky CT, Raftery AE. Bayesian information criterion for censored survival models. Biometrics. 2000; 56: 256-62.

7. Royston P, Sauerbrei W. Interaction of treatment with a continuous variable: simulation study of significance level for several methods of analysis. Stat Med. 2013; 32: 3788-803.

8. Wang W, Xu DZ, Li YF, Guan YX, Sun XW, Chen YB, et al. Tumor-ratio-metastasis staging system as an alternative to the 7th edition UICC TNM system in gastric cancer after D2 resection--results of a single-institution study of 1343 Chinese patients. Annals of oncology : official journal of the European Society for Medical Oncology / ESMO. 2011; 22: 2049-56.

9. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing

models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med. 1996; 15: 361-87.

10. Iasonos A, Schrag D, Raj GV, Panageas KS. How to build and interpret a nomogram for cancer prognosis. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2008; 26: 1364-70.

SUPPLEMENTARY TABLES AND FIGURES

Supplementary Table 1. Patient characteristics after multiple imputations for missing data in the SEER cohort (N=7,192)

Variable	Mean (SD)/N (%)
Race	
White	4,800 (66.7)
Black	944 (13.1)
Other	1,448 (20.1)
Tumor location	
Cardia	2,772 (38.5)
Upper one-third	743 (10.3)
Middle one-third	234 (3.3)
Lower one-third	3,443 (47.9)
Tumor size, cm	5.4 (3.0)
Tumor differentiation	
Poorly or undifferentiated	5,217 (72.5)
Well or moderately differentiated	1,975 (27.5)

SEER, Surveillance, Epidemiology, and End Results database; SD, standard deviation.

Supplementary Table 2. Patient characteristics pre- and post-weighted by inverse propensity score in the training set (N=5,206)

Variable	Pre-weighted cohort			Post-weighted cohort		
	Non-ART group (n=2,757)	ART group (n=2,449)	*P* value*	Non-ART group (n=2,757)	ART group (n=2,449)	*P* value*
Age			**<0.01**			0.21
Mean (SD)	69.3 (12.3)	60.4 (12.3)		64.9 (19.4)	64.5 (18.3)	
Gender			**<0.01**			0.40
Male	1,616 (58.6)	1,611(65.8)		1,724 (62.5)	1,559 (63.7)	
Female	1,141 (41.4)	838 (34.2)		1,033 (37.5)	890 (36.3)	
Tumor location			**<0.01**			0.96
Cardia	915 (33.2)	1,079 (44.1)		1,067 (38.7)	957 (39.1)	
Upper one-third	278 (10.1)	233 (9.5)		270 (9.8)	248 (10.1)	
Middle one-third	94 (3.4)	73 (3.0)		88 (3.2)	78 (3.2)	
Lower one-third	1,470 (53.3)	1,064 (43.4)		1,332 (48.3)	1,166 (47.6)	
Tumor size, cm			0.63			0.42
Mean (SD)	5.4 (2.8)	5.4 (3.1)		5.5 (4.4)	5.5 (4.1)	
Tumor differentiation			**<0.01**			0.49
Poorly or undifferentiated	1,903 (69.0)	1,865 (76.2)		2,009 (72.9)	1,805 (73.7)	
Moderately or well differentiated	854 (31.0)	584 (23.8)		748 (27.1)	644 (26.3)	
T stage			**<0.01**			0.92
T1	139 (5.0)	114 (4.7)		130 (4.7)	114 (4.7)	
T2	537 (19.5)	287 (11.7)		423 (15.3)	354 (14.5)	
T3	1,046 (37.9)	934 (38.1)		1,046 (37.9)	948 (38.7)	
T4a	748 (27.1)	865 (35.3)		866 (31.4)	768 (31.4)	
T4b	287 (10.4)	249 (10.2)		292 (10.6)	265 (10.8)	

MLN			**<0.01**			0.77
Mean	4.4 (6.5)	5.3 (6.4)		5.0 (9.5)	5.1 (8.9)	
NLN			0.56			0.85
Mean (SD)	10.8 (10.4)	10.7 (10.1)		10.7 (15.2)	10.8 (14.3)	
LNR			**<0.01**			0.91
Mean (SD)	0.30 (0.33)	0.35 (0.32)		0.34 (0.53)	0.33 (0.49)	
Race			**<0.01**			0.71
White	1,853 (67.2)	1,654 (67.5)		1,857 (67.3)	1,630 (66.6)	
Black	394 (14.3)	283 (11.6)		358 (13.0)	337 (13.7)	
Other	510 (18.5)	512 (20.9)		542 (19.7)	482 (19.7)	
Year of diagnosis			**0.01**			>0.99
2002	620 (22.5)	470 (19.2)		502 (20.5)	576 (20.9)	
2003	559 (20.3)	510 (20.8)		500 (20.4)	560 (20.3)	
2004	583 (21.1)	509 (20.8)		518 (21.2)	579 (21.0)	
2005	518 (18.8)	468 (19.1)		460 (18.8)	513 (18.6)	
2006	477 (17.3)	492 (20.1)		469 (19.2)	529 (19.2)	

ART, adjuvant radiotherapy; SD, standard deviation; MLN, metastatic lymph node; NLN, negative lymph node; LNR, lymph node ratio.
* Bold *P* values indicate statistical significance (i.e., $P<0.05$).

Supplementary Table 3. Patient characteristics pre- and post-weighted by inverse propensity score in the SEER validation set (N=1,986)

Variable	Pre-weighted cohort			Post-weighted cohort		
	Non-ART group (n=1,053)	ART group (n=953)	*P* value*	Non-ART group (n=1,053)	ART group (n=953)	*P* value*
Age			**<0.01**			0.53
Mean (SD)	68.08 (13.9)	60.7 (11.9)		64.3 (19.5)	63.9 (18.3)	
Gender			**<0.01**			0.15
Male	625 (59.4)	630 (67.5)		653 (61.9)	606 (65.1)	
Female	428 (40.6)	303 (32.5)		400 (38.1)	327 (34.9)	
Tumor location			**<0.01**			0.9
Cardia	356 (33.8)	422 (45.2)		425 (40.3)	390 (41.7)	
Upper one-third	129 (12.3)	103 (11.0)		124 (11.8)	105 (11.3)	
Middle one-third	38 (3.6)	29 (3.1)		33 (3.1)	31 (3.4)	
Lower one-third	530 (50.3)	379 (40.6)		471 (44.8)	407 (43.6)	
Tumor size, cm			0.36			0.29
Mean (SD)	5.3 (3.7)	5.5 (4.5)		5.4 (6.8)	5.6 (6.4)	
Tumor Differentiation			0.25			0.39
Poorly or undifferentiated	757 (71.9)	692 (74.2)		783 (74.4)	678 (72.6)	
Moderately or well differentiated	296 (28.1)	241 (25.8)		270 (25.6)	255 (27.4)	
T stage			**<0.01**			0.89
T1	64 (6.1)	51 (5.5)		62 (5.8)	58 (6.2)	
T2	193 (18.3)	113 (12.1)		158 (15.0)	133 (14.3)	
T3	459 (43.6)	406 (43.5)		454 (43.1)	408 (43.8)	
T4a	261 (24.8)	281 (30.1)		299 (28.4)	254 (27.2)	

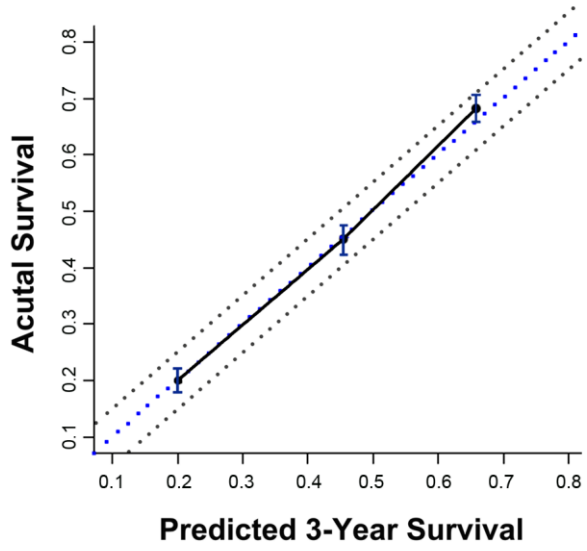
T4b	76 (7.2)	82 (8.8)		80 (7.6)	80 (8.6)	
MLN			0.10			0.20
Mean (SD)	4.5 (6.9)	5.0 (6.0)		4.8 (9.7)	5.2 (9.2)	
NLN			0.45			0.79
Mean (SD)	12.5 (11.3)	12.9 (11.2)		12.7 (16.2)	12.6 (15.3)	
LNR			0.05			0.47
Mean (SD)	0.27 (0.31)	0.30 (0.29)		0.29 (0.32)	0.30 (0.31)	
Race			**0.02**			0.24
White	660 (62.7)	633 (67.8)		679 (64.5)	619 (66.3)	
Black	142 (13.5)	125 (13.4)		134 (12.7)	130 (13.9)	
Other	251 (23.8)	175 (18.8)		240 (22.8)	184 (19.8)	
Year of diagnosis			0.35			0.84
2007	542 (51.5)	461 (49.4)		534 (50.7)	468 (50.2)	
2008	511 (48.5)	472 (50.6)		519 (49.3)	465 (49.8)	

Surveillance, Epidemiology, and End Results, SEER; ART, adjuvant radiotherapy; SD, standard deviation; MLN, metastatic lymph node; NLN, negative lymph node; LNR, lymph node ratio.

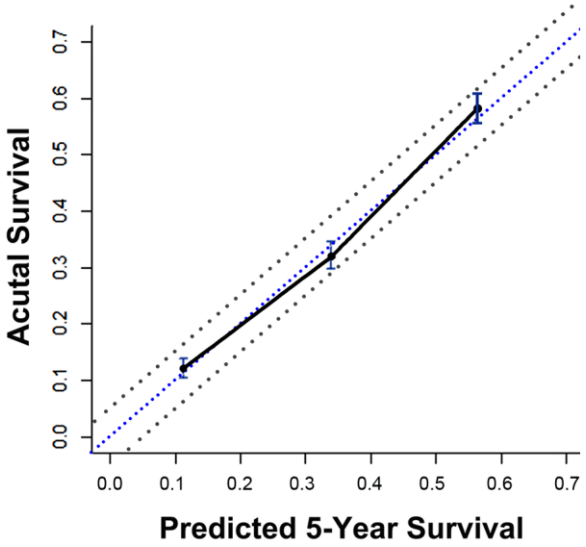
* Bold *P* values indicate statistical significance (i.e., $P<0.05$).

Supplementary Figure 1. Calibration plots of the nomograms for predicting 3- and 5-year overall survival in the training set (a and b) and the validation set (c and d). The diagonal blue dotted line represents the ideal reference line. The diagonal black dotted lines represent a 5% margin of error. The nomogram-predicted 3- and 5-year survival corresponded closely to the actual survival and was always within the 5% margin of error in both datasets.

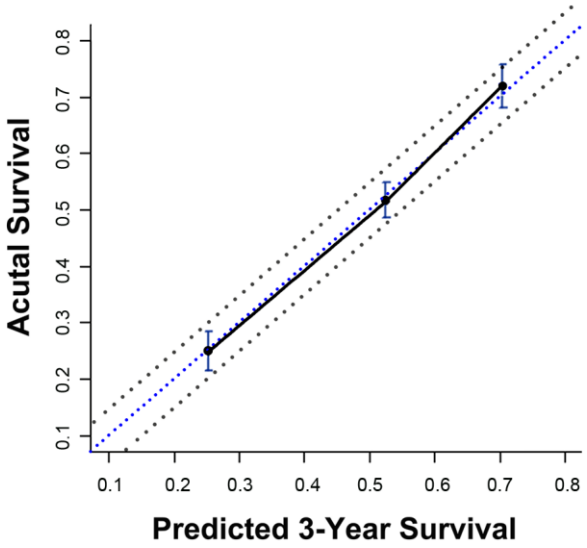
(a)


Acutal Survival
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Predicted 3-Year Survival

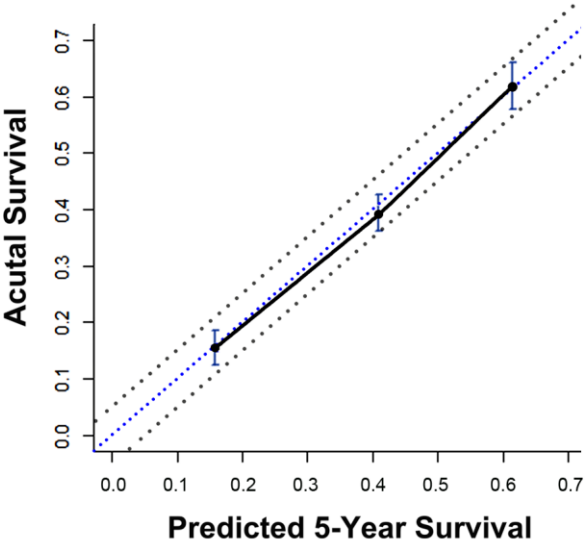
(b)


Acutal Survival
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Predicted 5-Year Survival

(c)


Acutal Survival
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Predicted 3-Year Survival

(d)


Acutal Survival
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
Predicted 5-Year Survival

Supplementary Figure 2. Web software to predict the 3- and 5-year overall survival probabilities with or without adjuvant radiotherapy among patients with resected gastric cancer. The nomogram-predicted 3- and 5-year overall survival probabilities with or without adjuvant radiotherapy can be calculated by entering clinicopathologic variables on the website.