

# The AUGIS Survival Predictor: Prediction of Long-Term and Conditional Survival After Esophagectomy Using Random Survival Forests

Saqib A. Rahman, MRCS,\*¶ Robert C. Walker, MRCS,\* Nick Maynard, FRCS,†  
Nigel Trudgill, MBBS,‡ Tom Crosby, FRCP,§ David A. Cromwell, PhD,¶ and  
Timothy J. Underwood, PhD,\*¶ on behalf of the NOGCA project team AUGIS

**Objective:** The aim of this study was to develop a predictive model for overall survival after esophagectomy using pre/postoperative clinical data and machine learning.

**Summary Background Data:** For patients with esophageal cancer, accurately predicting long-term survival after esophagectomy is challenging.

From the \*School of Cancer Sciences, Faculty of Medicine, University of Southampton, Southampton, UK; †Oxford University Hospitals NHS Trust, UK; ‡Sandwell and West Birmingham Hospitals NHS Trust, Birmingham, UK; §Velindre Cancer Center, Cardiff, UK; and ¶Clinical Effectiveness Unit, Royal College of Surgeons of England, London, UK.

✉tju@soton.ac.uk.

The authors report no conflicts of interest.

Patient and public involvement: Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research. Patient consent for publication was not required.

Ethics approval: The study is exempt from UK National Research Ethics Committee approval as it involved secondary analysis of an existing dataset of anonymized data. The National esophago-Gastric (OG) Cancer Audit has approval for processing health care information under Section 251 (reference number: ECC 1-06 (c)/2011) for all National Health Service (NHS) patients diagnosed with OG cancer in England and Wales. Data for this study are based on patient-level information collected by the NHS, as part of the care and support of patients with cancer.

S.A.R. is supported by a Royal College of Surgeons of England Research Fellowship and a British Association of Surgical Oncologists Research Project Grant.

T.J.U. is supported by a Cancer Research UK and Royal College of Surgeons of England Advanced Clinician Scientist Fellowship, ID: A23924.

This project has been supported by the Association of Upper Gastrointestinal Surgery (AUGIS), Heartburn Cancer UK, and The Royal College of Surgeons of England Surgical Specialty Lead Programme.

Disclaimer: Neither HQIP nor the funders had any involvement in the study design; in the collection, analysis and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

NOGCA Project Team: Clinical Effectiveness Unit: D.C., M.H.P., H.W.

Clinical leads: N.M., T.C., N.T.

NHS Digital: J.G. and R.N.

Funding: This study was undertaken as part of the work by the National esophago-Gastric (OG) Cancer Audit. The Audit is commissioned by the Healthcare Quality Improvement Partnership (HQIP) as part of the National Clinical Audit and Patient Outcomes Programme and funded by NHS England and the Welsh Government ([www.hqip.org.uk/national-programmes](http://www.hqip.org.uk/national-programmes)). The authors had full independence from the Healthcare Quality Improvement Partnership. The aim of National esophago-Gastric Cancer Audit is to evaluate the care of patients with OG cancer in England and Wales, and support NHS providers to improve the quality of hospital care for these patients. More information can be found at: [www.nogca.org.uk](http://www.nogca.org.uk).

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website, [www.annalsofsurgery.com](http://www.annalsofsurgery.com).

This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2021 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 0003-4932/23/27702-0267

DOI: 10.1097/SLA.0000000000004794

This study investigated survival prediction after esophagectomy using a Random Survival Forest (RSF) model derived from routine data from a large, well-curated, national dataset.

**Methods:** Patients diagnosed with esophageal adenocarcinoma or squamous cell carcinoma between 2012 and 2018 in England and Wales who underwent an esophagectomy were included. Prediction models for overall survival were developed using the RSF method and Cox regression from 41 patient and disease characteristics. Calibration and discrimination (time-dependent area under the curve) were validated internally using bootstrap resampling.

**Results:** The study analyzed 6399 patients, with 2625 deaths during follow-up. Median follow-up was 41 months. Overall survival was 47.1% at 5 years. The final RSF model included 14 variables and had excellent discrimination with a 5-year time-dependent area under the receiver operator curve of 83.9% [95% confidence interval (CI) 82.6%–84.9%], compared to 82.3% (95% CI 81.1%–83.3%) for the Cox model. The most important variables were lymph node involvement, pT stage, circumferential resection margin involvement (tumor at < 1 mm from cut edge) and age. There was a wide range of survival estimates even within TNM staging groups, with quintiles of prediction within Stage 3b ranging from 12.2% to 44.7% survival at 5 years.

**Conclusions:** An RSF model for long-term survival after esophagectomy exhibited excellent discrimination and well-calibrated predictions. At a patient level, it provides more accuracy than TNM staging alone and could help in the delivery of tailored treatment and follow-up.

**Keywords:** esophageal cancer, esophagus, machine learning, prognostic model

(*Ann Surg* 2023;277:267–274)

Esophagectomy for cancer is a highly morbid operation from which patients frequently take > 18 months to recover.<sup>1–3</sup> Long-term prognosis for patients also remains poor, with 5-year survival estimated to be < 50%.<sup>4</sup>

Presently, clinicians have a limited number of tools to identify patients with esophageal cancer who are likely to respond well to surgery and those who may not. TNM staging is widely used for patient stratification, but the classification is based on largely historic data (patients treated in 1980 s–2000 s).<sup>5</sup> In addition, staging groups remain coarse, even with the introduction of post-neoadjuvant staging (ie, ypTNM) in TNM 8.<sup>6</sup> Important characteristics that are readily available and routinely collected (such as circumferential resection margin) are not considered for the sake of simplicity, leading to a range of survival outcomes for patients within the same stage groups. This makes application at the patient level inaccurate.

The delivery of personalized long-term survival estimates after treatment for esophageal cancer is challenging. In addition to informing patients, reliable survival figures would enable the

identification of high-risk individuals or groups in whom enhanced surveillance or treatment intensification (with traditional or novel agents such as immunotherapy) could be considered, or conversely patients where de-escalation would be the preferred option.

Prognostic models can address these limitations by combining multiple risk factors, although none have entered widespread use among surgeons or oncologists treating esophageal cancer.<sup>7,8</sup> Models based on machine learning (ML) techniques may produce more accurate predictions than models built using traditional statistical methods (eg, logistic/cox regression).<sup>9,10</sup> In particular, Random Survival Forest (RSF) models have produced promising results<sup>11–13</sup> in various settings, and in esophageal cancer were used to derive the AJCC TNM 7<sup>th</sup> and 8<sup>th</sup> edition staging manuals,<sup>5,6</sup> and to quantify the benefits of optimizing treatment.<sup>14</sup> RSF is a machine-learning method that, when developed to predict survival, builds many decision trees with log-rank test based split points to identify different survival trajectories, with the predicted probability for an individual being derived as the average prediction across all of the trees.

The aim of this study was to derive and validate a prognostic model based on RSF methods for long-term survival after esophagectomy for cancer, and to compare its performance to a model developed using a common statistical approach (Cox regression), using a population-based dataset from England and Wales.

## METHODS

### Study Cohort

The study used a linked dataset prepared by the National Esophago-Gastric Cancer Audit (NOGCA), a national clinical audit of patients undergoing treatment for cancer of the esophagus or stomach in England and Wales.<sup>15</sup> The audit was commissioned by the Healthcare Quality Improvement Partnership (HQIP) and funded by NHS England and the Welsh government. Patients were eligible for inclusion in the audit if they had a histological diagnosis of epithelial cancer, with the first patients being registered in April 2012. The audit collects a dataset that covers the care pathway from diagnosis to the end of initial treatment and links these patient records with information from other national health care datasets, including the National Cancer Registration and Analysis Service (NCRAS, see<sup>15</sup> for more details). Data collection was approved by the Confidentiality Advisory Group under section 251 of the NHS Act 2006.

### Ethics Approval

The study is exempt from UK National Research Ethics Committee approval as it involved secondary analysis of an existing dataset of anonymized data. The National Esophago-gastric (OG) Cancer Audit has approval for processing health care information under Section 251 (reference number: ECC 1–06 (c)/2011) for all National Health Service (NHS) patients diagnosed with OG cancer in England and Wales. Data for this study are based on patient-level information collected by the NHS, as part of the care and support of patients with cancer.

The study cohort included patients diagnosed with adenocarcinoma or squamous cell carcinoma of either the esophagus or gastroesophageal junction (Siewert I–II) between April 1, 2012 and March 31, 2018 who underwent a planned curative esophagectomy. The study excluded patients who died in hospital before discharge, had confirmed metastatic disease on

postoperative histology, or had an inadequate lymphadenectomy (< 15 lymph nodes),<sup>16</sup> in whom interpretation of lymph node status would be biased. Supplementary Figures S.1, <http://links.lww.com/SLA/C964> and S.2, <http://links.lww.com/SLA/C964> details the patient exclusions and assumptions to derive the final sample size (n = 6399).

The primary outcome was overall survival from the date of discharge following surgery. Survival was confirmed by linking the audit records with records from the Office for National Statistics death register. Median duration of follow-up was 41 months (interquartile range 24–59).

### Variable Definition

The audit data contained 41 variables that were routinely measured in clinical practice, were beyond the control of the provider, had > 50% completeness, and were clinically relevant to survival, listed in Table S.1, <http://links.lww.com/SLA/C964>. The dataset contained patient characteristics, disease information, details of treatment received, postoperative complications, and tissue pathology. Circumferential resection margins were considered involved if there was tumor at < 1 mm from the cut edge and longitudinal resection margins were considered involved if tumor was found at the cut edge, in line with Royal College of Pathologists Guidelines.<sup>17</sup> In patients undergoing neoadjuvant therapy, treatment was specified as “complete” if it was completed as prescribed or “not complete” (due to disease progression, treatment toxicity, technical problems, or patient choice). Malignant esophageal and gastric surgery is centralized in England and Wales and undertaken solely by dedicated teams. We therefore defined annual hospital volume as average number of major upper gastrointestinal resections (esophagectomy/major gastrectomy) per year, in line with NHS commissioning guidelines.<sup>18</sup> Staging was conducted using the 8<sup>th</sup> edition of the AJCC TNM staging manual.

Among the 41 variables considered for inclusion, five had missing values for > 5% of patients: completion of neoadjuvant treatment (19.9%), return to theater (15.8%), grade of differentiation (7.0%), cT stage (5.9%), and surgical approach (5.7%). Missing data were assumed to be missing at random and were addressed using multiple imputation by chained equations with 10 imputations.<sup>19</sup>

### Model Development

The study aimed to develop a model using a subset of variables so that data collection would be straightforward, and the model easy to use in clinical settings. To select core variables, we used permutation based Random Forest variable importance (VIMP)<sup>11</sup> with bootstrapped confidence intervals. Variables with a  $P < 0.01$  for  $VIMP > 0$  were included in the final model (Table S.2, <http://link-s.lww.com/SLA/C964>). Pre-treatment histology (ie, adenocarcinoma or squamous cell carcinoma) was also included to improve the face validity of the model. The final model was trained using 14 variables: age, sex, cT, cN, site of tumor, pre-treatment histology, neoadjuvant treatment, completion of neoadjuvant treatment, pT/ypT, number of positive lymph nodes, circumferential and longitudinal margin involvement, grade of differentiation, and presence/absence of surgical complications. The RSF hyperparameters (ie, number of trees, number of variables per tree, and minimum node size) were optimized by grid search. Final predictions were combined across the imputed data.<sup>20–22</sup>

A cox regression model was also developed using the same set of variables. Not all relationships between survival and continuous variables were linear, and a square root

transformation was adopted for positive lymph nodes, whereas age was included as a restricted cubic spline.

### Assessment of Model Performance

Model performance was quantified by discrimination and calibration. Discrimination was assessed using the time-dependent area under the receiver operator curve (tAUC).<sup>23</sup> Here this represents the proportion of random pairs of patients (1 alive at time point “t” and 1 dead before this) where the model gives the patient who is alive a higher probability of survival than the patient who is dead. It can be considered analogous to the standard AUC in a binary regression model, extended to survival by weighting of censored patients,<sup>24</sup> and has advantages over the C-statistic measure of performance.<sup>25</sup> Assessment of calibration was conducted visually for 5 patient subgroups of

increasing risk (ie, patients were grouped by quintiles of predicted risk of mortality at 5 years). In addition, we calculated the integrated brier score.<sup>26,27</sup> A score closer to 0 indicates better accuracy of predictions.

Finally, the relative performance of the 2 models was compared using decision curve analysis.<sup>28</sup> This method is based on evaluating the “net-benefit” of model predictions across of range of possible decision thresholds that reflect how a patient might weigh the risk of harm associated with a false-positive result (compared with a true positive result). Models with a better performance have a greater net benefit across all thresholds of probability.

Data analysis was conducted in R 3.5.3.<sup>29</sup> The RSF was trained using the packages Ranger<sup>30</sup> and RF-SRC.<sup>31</sup> This study was conducted to comply with the AJCC prognostic model<sup>32</sup> and

**TABLE 1.** Background Characteristics of Patients Who Underwent An Esophagectomy Between April 2012 and March 2018

Characteristic	N = 6399	Median Survival	Characteristic	N = 6399	Median Survival
Sex			Anastomotic leak		
Male	5045 (78.8)	47	No	5923 (92.6)	54
Female	1354 (21.2)	72	Yes	445 (7.0)	40
Age			Unknown	31 (0.5)	NR
0–40	64 (1.0)	NR	Any complication		
41–50	405 (6.3)	68	No	3810 (59.5)	55
51–60	1397 (21.8)	59	Yes	2558 (40.0)	49
61–70	2615 (40.9)	60	Unknown	31 (0.5)	
71–80	1787 (27.9)	42	Involved longitudinal margin		
81 +	131 (2.0)	29	No	6188 (96.7)	55
Site of tumor			Yes	211 (3.3)	19
Upper/mid esophagus	792 (12.4)	61	Involved circumferential margin		
Lower esophagus	3795 (59.3)	52	No	4617 (72.2)	77
GEJ (S1–2)	1812 (28.3)	53	Yes	1534 (24.0)	21
Histopathology			Unknown	248 (3.9)	57
Adenocarcinoma	5540 (86.6)	51	pT/ypT		
SCC	859 (13.4)	68	T0/is	524 (8.2)	NR
cT			T1	1201 (18.8)	NR
T0/is/1	467 (7.3)	NR	T2	836 (13.1)	NR
T2	1294 (20.2)	67	T3	3549 (55.5)	30
T3	3979 (62.2)	38	T4	289 (4.5)	13
T4	284 (4.4)	36	Lymph nodes examined	26 [15–130]	
Unknown	375 (5.9)	NR	pN/ypN		
cN			N0	2994 (46.8)	NR
N0	2551 (39.9)	76	N1	1414 (22.1)	46
N1	2547 (39.8)	41	N2	1133 (17.7)	22
N2	938 (14.7)	32	N3	858 (13.4)	14
N3	159 (2.5)	28	Grade		
Unknown	204 (3.2)	47	G1 (well)	226 (3.5)	NR
cM			G2 (moderate)	2331 (36.4)	60
M0	6151 (96.1)	54	G3/4 (poor/anaplastic)	2697 (42.1)	38
M1	44 (0.7)	26	GX (unable to determine)	695 (10.9)	66
Unknown	204 (3.2)	47	Unknown	450 (7.0)	72
ASA			NAT		
1	892 (13.9)	64	Chemotherapy	3976 (62.1)	42
2	3745 (58.5)	53	Chemoradiotherapy	450 (7.0)	NR
3	1726 (27.0)	42	None	1973 (30.8)	66
4	36 (0.6)	43	Completion of NAT		
Approach			Completed	2981 (46.6)	51
Open	3357 (52.5)	51	Not completed	282 (4.4)	32
Hybrid	1931 (30.2)	51	Not applicable	1861 (29.1)	67
MIO	748 (11.7)	NR	Unknown	1275 (19.9)	42
Unknown	363 (5.7)	41	Annual hospital volume		
			1–30	504 (7.9)	59
			31–60	3351 (52.4)	55
			> 60	2544 (39.8)	48

Data given as absolute number (percentage) and median (Range). MIO indicates minimally invasive esophagectomy; NAT, neoadjuvant treatment; NR, median survival not reached; SCC, squamous cell carcinoma.

TRIPOD<sup>33</sup> criteria, a compliance checklist is provided in Table S.3, <http://links.lww.com/SLA/C964>. Complete R code to reproduce the analysis is available on request. More extensive methodology and instructions to perform external validation are provided in the supplementary materials. All performance metrics were validated internally by the 0.632 estimator<sup>34</sup> in 1000 replications of the bootstrap with replacement.

### RESULTS

The study included 6399 patients with esophageal cancer who underwent an esophagectomy between April 2012 and March 2018. Table 1 summarizes the characteristics of patients and their treatment. The median age at diagnosis was 66 years and only 1 in 5 were women. Tumors were predominantly adenocarcinoma (87%) and about 3 in 10 were classified as GEJ-Siewert I-II. There were 2625 recorded deaths, and the median survival was 53 months. Survival at 1, 3, and 5 years was 83.7%, 57.1%, and 47.1% respectively (Fig. 1). Differences in survival stratified by stage according to if patients received neoadjuvant treatment (ie, ypTNM) or surgery alone (ie, pTNM) are shown and discussed in supplementary Figure S.9, <http://links.lww.com/SLA/C964>.

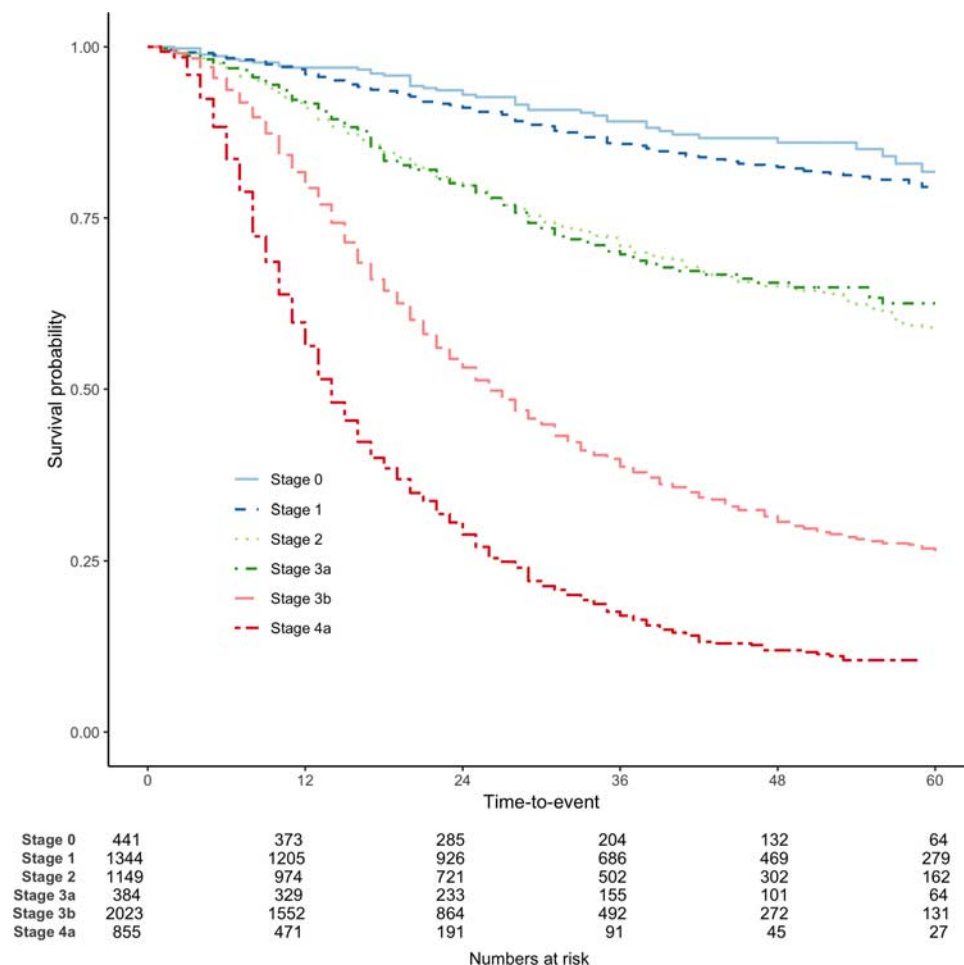
A total of 13 variables were identified as important to include in the final model in addition to histological diagnosis. The RSF variable importance measure indicated the number of

lymph nodes as the most important single risk factor for worse prognosis followed by pT/ypT stage (see partial dependence plots, supplementary figures S.3/S.4, <http://links.lww.com/SLA/C964>).

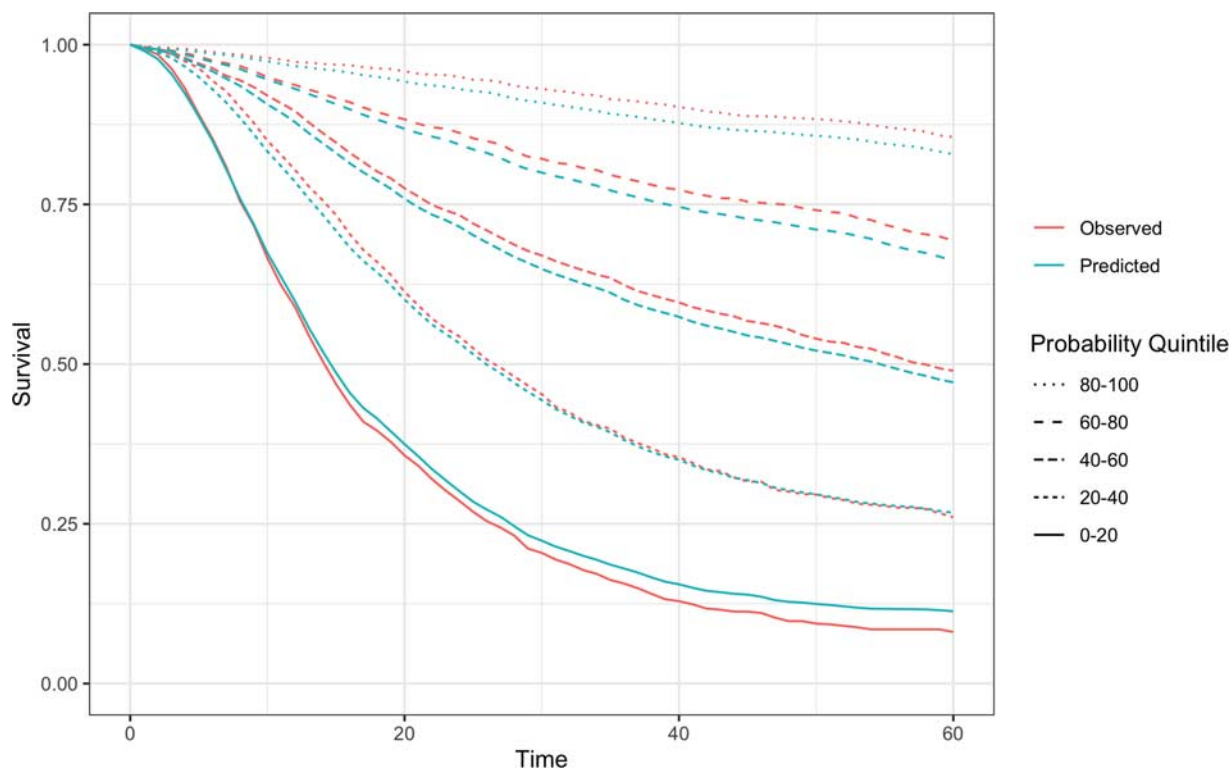
### Model Performance: Internal Validation

The RSF model demonstrated excellent discrimination, with a bootstrapped tAUC at 60 months of 83.9% (95% CI 82.6%–84.9%), which was similar at other time points (Figure S.6, <http://links.lww.com/SLA/C964>). This was better than the Cox regression model (coefficients of which are given in Table S.4, <http://links.lww.com/SLA/C964>), which had a bootstrapped tAUC of 82.3% (95% CI 81.1%–83.3%) and TNM stage alone (tAUC 74.5%). Figure 2 shows the agreement between the RSF model predicted and observed survival times for patients grouped according to quintile of prediction and in both models, calibration was visually good throughout these groups. The integrated brier scores for the RSF model was superior to the cox regression at 0.136 (95% CI 0.134–0.138) and 0.141 (0.139–0.143), respectively. Decision curve analysis also showed a greater net benefit for the RSF over Cox regression model (Figure S.7, <http://links.lww.com/SLA/C964>) or using TNM alone (Figure S.8, <http://links.lww.com/SLA/C964>).

There were a broad range of predictions yielded even within p/ ypTNM staging groups, with the lowest risk quintile of Stage 3b patients having a predicted 5-year survival of 44.7%



**FIGURE 1.** Survival of patients who underwent an esophagectomy between April 2012 and March 2018, stratified by TNM stage.



**FIGURE 2.** Calibration of predictions from RSF model. Patients grouped into quintiles according to predicted survival at 60 months post-surgery.

compared to 12.2% for the highest risk quintile. Moreover, there is a subgroup of early-stage disease (TNM stage 0–1), who would generally be considered to be cured, who had a relatively poorer survival of only 64.7% at 5 years (Fig. 3) and overlap of quintiles between staging groups.

Figure 4 gives an overview of mean predicted 5-year survival for combinations of the most important variables (Lymph node status, T-stage, circumferential resection margin involvement, and age at diagnosis). Age at diagnosis is most influential with early stage (T0–2, N0–1) disease; however, its importance diminishes with increasing T/N-stage. Examples of how the model might be used are given in the supplementary materials (Figure S.10, <http://link-s.lww.com/SLA/C964>)

## DISCUSSION

Accurate predictions of long-term survival following surgery for esophageal cancer may help clinicians and patients. This study has demonstrated that an RSF model can discriminate between patients with different long-term prognoses using a small number of routinely collected variables. The model showed very good calibration and discrimination on internal validation, and exceeded that achieved using Cox regression analysis. The model is applicable to patients who have undergone a planned curative esophagectomy for adenocarcinoma or squamous cell carcinoma of the esophagus, who had an adequate lymphadenectomy and survived to discharge from hospital.

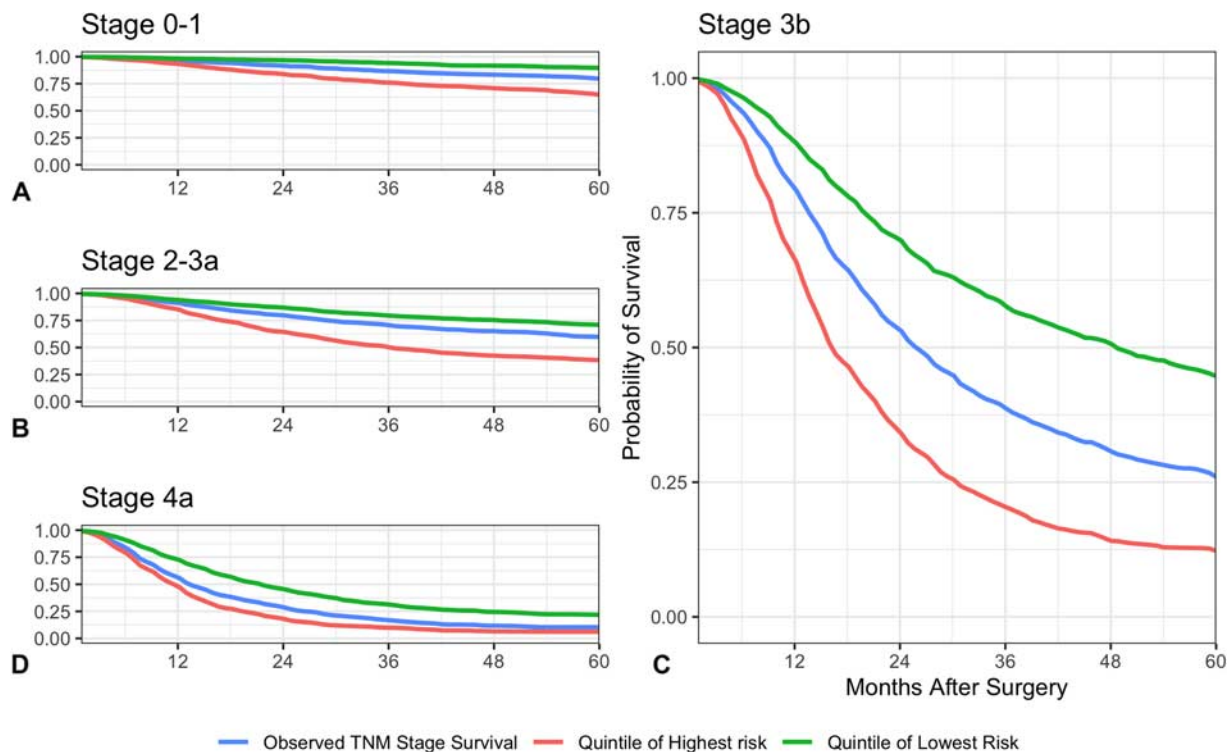
At present, information given to patients after surgery about their long-term survival is limited and is largely based on TNM staging. This can mean the information provided to

patients can be vague, such as “50% survive to 5 years.” Decisions on whom to offer adjuvant treatment or consider for entry into trials may involve more criteria than TNM staging, but the relationship between these criteria and survival may be uncertain.

The model described here provides a more precise prediction of prognosis for an individual patient than TNM staging alone, and this will be valuable in postoperative discussions with patients. This increased accuracy has several benefits. In a research setting, it is key for establishing the efficacy of treatment. In clinical practice, it supports selecting the right patients for the right treatments, particularly with the emergence of novel therapies (eg, Immune checkpoint inhibitors<sup>35</sup>). Further research on how best to communicate predicted survival to patients is required; even in early-stage disease, desire for detail of prognosis is highly variable,<sup>36</sup> and the effective use of decision aids is challenging.<sup>37</sup>

The model compares favorably to those published previously. Cox proportional hazard models using a variety of predictors have reported C-index/tAUCs between 0.61 and 0.70.<sup>38–40</sup> In comparison the C-index of our model was 0.76 (0.75–0.78) and the 5-year tAUC 0.84 (0.83–0.85). It also is more broadly applicable and includes patients with all modalities of neoadjuvant treatment.

The variables found to be most influential are consistent with clinical experience and the findings from other studies, with lymph node involvement being widely recognized as the most influential determinant of long-term survival in esophageal cancer. Both clinical T stage and clinical N stage were found to be important independent of their pathological equivalents. There is some logic to support including both clinical and



**FIGURE 3.** Range of predictions by p/ypTNM stage. A, Stage 0–1. B, Stage 2–3a. C, Stage 3b. D, Stage 4a. Patients were grouped into quintiles by predicted survival at 60 months, with the highest and lowest groups shown.

pathological variables, in that changing variables within patients may indicate the impact of neoadjuvant treatment (although this is limited by the relatively decreased accuracy of clinical staging). This is supported by recent studies which have shown that downstaging after neoadjuvant treatment improves absolute survival independent of the ypTNM stage.<sup>41</sup> Completion of neoadjuvant treatment was included which is biologically sensible, and important in the context of the increasing use of potentially more toxic regimens such as FLOT.<sup>4</sup>

The main strength of this study is the large sample size from a national population. The case ascertainment of esophagectomies exceeds 90% in the national audit, and the dataset was representative of patients within England and Wales who underwent curative surgery. Another strength is the linkage of audit records with ONS mortality data which enabled complete follow-up.

There are a number of limitations to the approaches taken in this study. Despite being more accurate than TNM staging at the individual patient level in the postoperative setting, no attempt has been made to develop a pre-treatment predictive model and cTNM remains the criterion standard in this domain. The NOGCA lacks several data items known to influence survival such as tumor regression grade<sup>42</sup> and lymphovascular invasion.<sup>43</sup> Additionally length of tumor<sup>44</sup> and BMI<sup>45</sup> could be considered, but were only available in recent years and therefore had too many missing values. There was also no clear information on what adjuvant treatment this patient cohort had received in addition to their neoadjuvant/ perioperative treatment.

Involvement of circumferential margin was defined according to Royal College of Pathologists criteria, that is, <1 mm from cut edge is involved. Throughout much of the rest of

the world the American College of Pathologist guidelines are used,<sup>46</sup> that is, involved if tumor at cut edge. There is considerable debate about the most appropriate measure,<sup>47</sup> and this model requires validation if it is to be used with this definition. It was not possible to use T stage subdivided into “a” or “b” because not all patients were recorded with this information. Consequently, the analysis used the base T stage only. Patients treated solely with endoscopic techniques (mucosal resection or submucosal dissection) who did not require surgery were also excluded and it is not appropriate to use the model in this patient group.

The Esophageal Complications Consensus Group—ECCG<sup>48</sup> has recently specified and defined a core set of complications for esophagectomy which has been adopted worldwide.<sup>49</sup> The NOGCA relies on reporting from local cancer centers and pragmatically uses a limited set of complications with broader definitions. In this study, the reported rate of complications was 40.0%, which is significantly less than the figures from the ECCG data (59%). This is likely to reflect the varying definitions and under reporting rather than a truly lower rate, which may explain the low overall importance of complications and absence of specific complications (eg, anastomotic leak) in the model. The ECCG classification of complications has recently been adopted into the national Cancer Outcomes and Services Dataset used in cancer registration within England, so more accurate complication data will be available in future iterations of the model.

## CONCLUSIONS

Using a large, nation-wide, contemporaneous clinical dataset, this study has demonstrated the ability of a Random

		Lymph Node Status									
		N0		N1		N2		N3			
		CRM		CRM		CRM		CRM			
		-ve	+ve	-ve	+ve	-ve	+ve	-ve	+ve		
p/ypT Stage	T0	Age	18-50	82.6%	65.2%	74.6%	54.6%	33.9%	31.7%	30.9%	25.6%
			51-60	83.3%	65.5%	74.8%	54.7%	33.6%	31.6%	30.4%	25.5%
			61-70	85.0%	66.2%	74.1%	54.3%	33.4%	29.9%	30.6%	23.7%
			71+	83.0%	63.2%	70.4%	52.5%	32.8%	30.6%	29.4%	22.8%
	T1	Age	18-50	73.6%	64.5%	69.9%	55.2%	34.1%	32.0%	31.9%	26.3%
			51-60	74.3%	64.8%	70.1%	55.2%	33.7%	31.9%	31.4%	26.2%
			61-70	76.8%	65.9%	69.7%	54.9%	33.6%	30.1%	31.7%	24.3%
			71+	76.6%	63.4%	66.7%	53.1%	32.9%	30.9%	30.4%	23.4%
	T2	Age	18-50	81.0%	67.3%	73.7%	56.4%	36.0%	32.6%	33.4%	26.6%
			51-60	81.2%	67.4%	73.7%	56.4%	35.6%	32.5%	32.9%	26.5%
			61-70	81.7%	67.6%	72.8%	56.0%	35.2%	30.5%	32.9%	24.4%
			71+	75.6%	62.6%	69.4%	54.1%	34.0%	31.0%	31.1%	23.3%
	T3	Age	18-50	74.0%	58.4%	59.6%	42.5%	33.6%	27.3%	25.1%	20.6%
			51-60	74.5%	58.4%	59.3%	42.3%	33.1%	27.1%	24.4%	20.4%
			61-70	68.7%	57.8%	54.8%	41.1%	33.0%	24.6%	26.0%	18.0%
			71+	62.3%	48.5%	50.3%	38.4%	30.0%	25.0%	23.5%	16.6%
	T4	Age	18-50	36.1%	22.0%	31.0%	18.3%	13.7%	13.1%	11.0%	11.0%
			51-60	36.2%	22.1%	31.0%	18.3%	13.6%	13.1%	11.0%	11.0%
			61-70	33.9%	21.3%	28.9%	17.5%	12.9%	12.3%	10.6%	10.3%
			71+	31.4%	19.6%	26.8%	16.7%	11.9%	11.4%	9.6%	9.2%
Key: Survival at 5 years		100%	50%	0%	CRM = Circumferential Resection Margin						

**FIGURE 4.** Predicted 5-year survival for a given combination of selected variables. Colors represent differing prognosis, with green more favourable and red less favourable.

Survival Forest model to provide accurate predictions of long-term survival after surgery for esophageal cancer. A key benefit of the model is its performance in identifying patients with the same disease stage who have diverging 5-year survival. For example within Stage 3b, the largest group with 2023 patients, the model identifies a low-risk quintile of patients with a predicted 5-year survival > 3 times the highest risk quintile (44.7% vs 12.2%). These groups will likely benefit from different post-operative monitoring and/or treatment strategies. A similar pattern is seen with stage 4a disease (21.7% vs 6.1% 5-year survival), suggesting that there is a subgroup even in the most advanced (nonmetastatic disease) who might be well served by targeted intervention.

The RSF model described in this article is available at <https://uoscancer.shinyapps.io/AugisSurv/> and could be a valuable prognostication tool for patients, surgeons, and oncologists. In the future, it may also be useful to guide adjuvant treatment. External validation of this tool in other health care systems would be of benefit to confirm its performance.

## REFERENCES

- Zhang Y, Yang X, Geng D, et al. The change of health-related quality of life after minimally invasive esophagectomy for esophageal cancer: a metaanalysis. *World J Surg Oncol*. 2018;16:1–13.
- Derogar M, Lagergren P. Health-related quality of life among 5-year survivors of esophageal cancer surgery: a prospective population-based study. *J Clin Oncol*. 2012;30:413–418.
- Geeraerts M, Silva Corten LC, van Det M, et al. Insights in work rehabilitation after minimally invasive esophagectomy. *Surg Endosc*. 2019;33:3457–3463.
- Al-Batran S-E, Homann N, Pauligk C, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet*. 2019;393:1948–1957.
- Rice TW, Rusch VW, Ishwaran H, et al. Cancer of the esophagus and esophagogastric junction: data-driven staging for the seventh edition of the American Joint Committee on Cancer/International Union Against Cancer Staging Manuals. *Cancer*. 2010;116:3763–3773.
- Rice TW, Patil DT, Blackstone EH. 8th edition AJCC/UICC staging of cancers of the esophagus and esophagogastric junction: application to clinical practice. *Ann Cardiothorac Surg*. 2017;6:119–130.

7. Gupta V, Coburn N, Kidane B, et al. Survival prediction tools for esophageal and gastroesophageal junction cancer: a systematic review. *J Thorac Cardiovasc Surg.* 2018;156:847–856.
8. van den Boorn HG, Engelhardt EG, van Kleef J, et al. Prediction models for patients with esophageal or gastric cancer: a systematic review and meta-analysis. *PLoS One.* 2018;13:e0192310.
9. Caruana R. An Empirical Comparison of Supervised Learning Algorithms. In: *Proceedings of the 23rd International Conference on Machine Learning.* Pittsburgh; 2006
10. Fernández-Delgado M, Cernadas E, Barro S, et al. Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res.* 2014;15:3133–3181.
11. Ishwaran H, Kogalur UB, Blackstone EH, et al. Random survival forests. *Ann Appl Stat.* 2008;2:841–860.
12. Hu C, Steingrimsson JA. Personalized risk prediction in clinical oncology research: applications and practical issues using survival trees and random forests. *J Biopharm Stat.* 2018;28:333–349.
13. Dietrich S, Floegel A, Troll M, et al. Random Survival Forest in practice: a method for modelling complex metabolomics data in time to event analysis. *Int J Epidemiol.* 2016;45:1406–1420.
14. Rice TW, Lu M, Ishwaran H, et al. Precision surgical therapy for adenocarcinoma of the esophagus and esophagogastric junction. *J Thorac Oncol.* 2019;14:2164–2175.
15. Cromwell D, Wahedly H, Park MH, et al. National Oesophago-Gastric Cancer Audit 2019. 2019. Available at: [https://www.hqip.org.uk/wp-content/uploads/2019/12/REF150\\_NOGCA\\_2019-Annual-Report-FINAL\\_v2\\_2019.12.18.pdf](https://www.hqip.org.uk/wp-content/uploads/2019/12/REF150_NOGCA_2019-Annual-Report-FINAL_v2_2019.12.18.pdf).
16. Allum WH, Blazeby JM, Griffin SM, et al. Guidelines for the management of oesophageal and gastric cancer. *Gut.* 2011;60:1449–1472.
17. RCPATH Cancer Services Working Group. Dataset for the Histopathological Reporting of Oesophageal Carcinoma (2nd Edition); 2013.
18. NHS England. 2013/14 NHS Oesophageal and Gastric Cancer Commissioning Guidelines; 2013.
19. van Buuren S, Groothuis-Oudshoorn K. MICE: multivariate imputation by chained equations in R. *J Stat Softw.* 2011;45:1–67.
20. Marshall A, Altman DG, Holder RL, et al. Combining estimates of interest in prognostic modelling studies after multiple imputation: current practice and guidelines. *BMC Med Res Methodol.* 2009;9:1–8.
21. Wood AM, Royston P, White IR. The estimation and use of predictions for the assessment of model performance using large samples with multiply imputed data. *Biometrical J.* 2015;57:614–632.
22. Hosmer D, Lemeshow S. In: *Applied Survival Analysis: Regression Modeling of Time to Event Data.* 1st ed., New York: John Wiley & Sons Inc.; 1999.
23. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol.* 2017;17:1–19.
24. Blanche P, Dartigues J-F, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med.* 2013;32:5381–5397.
25. Blanche P, Kattan MW, Gerds TA. The c-index is not proper for the evaluation of 't'-year predicted risks. *Biostatistics.* 2018;20:347–357.
26. Graf E, Schmoor C, Sauerbrei W, et al. Assessment and comparison of prognostic classification schemes for survival data. *Stat Med.* 1999;18:2529–2545.
27. Kronek LP, Reddy A. Logical analysis of survival data: Prognostic survival models by detecting high-degree interactions in right-censored data. *Bioinformatics.* 2008;24:248–253.
28. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Mak.* 2006;26:565–574.
29. R Core Team. R: A Language and Environment for Statistical Computing. 2019.
30. Wright MN, Ziegler A. Ranger: a fast implementation of random forests for high dimensional data in C++ and R. *J Stat Softw.* 2017;77:1–17.
31. Ishwaran H, Kogalur UB. Fast Unified Random Forests for Survival, Regression, and Classification (RF-SRC), Rpackage version 291 2019. Available at: <https://cranr-project.org/web/packages/randomForestSRC/randomForestSRC.pdf> Accessed November 27, 2019.
32. Moons KGM, Weiser MR, Lu Y, et al. American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine. *CA Cancer J Clin.* 2016;66:370–374.
33. Collins GS, Reitsma JB, Altman DG, et al. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD): The TRIPOD Statement. *Ann Intern Med.* 2015;162:55–63.
34. Efron B. Estimating the error rate of a prediction rule: Improvement on cross-validation. *J Am Stat Assoc.* 1983;78:316–331.
35. Kato K, Cho BC, Takahashi M, et al. Nivolumab versus chemotherapy in patients with advanced oesophageal squamous cell carcinoma refractory or intolerant to previous chemotherapy (ATTRACTION-3): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol.* 2019;20:1506–1517.
36. Mühlbauer V, Berger-Höger B, Albrecht M, et al. Communicating prognosis to women with early breast cancer—overview of prediction tools and the development and pilot testing of a decision aid. *BMC Health Serv Res.* 2019;19:1–15.
37. Kopecky KE, Urbach D, Schwarze ML. Risk calculators and decision aids are not enough for shared decision making. *JAMA Surg.* 2019;154:3–4.
38. Gabriel E, Attwood K, Shah R, et al. Novel calculator to estimate overall survival benefit from neoadjuvant chemoradiation in patients with esophageal adenocarcinoma. *J Am Coll Surg.* 2017;224:884–894e1.
39. Goense L, van Rossum PSN, Xi M, et al. Preoperative nomogram to risk stratify patients for the benefit of trimodality therapy in esophageal adenocarcinoma. *Ann Surg Oncol.* 2018;25:1598–1607.
40. Hagens ERC, Feenstra ML, Eshuis WJ, et al. Conditional survival after neoadjuvant chemoradiotherapy and surgery for oesophageal cancer. *Br J Surg.* 2017;104:1816–1828.
41. Kamarajah SK, Navidi M, Wahed S, et al. Significance of neoadjuvant downstaging in carcinoma of the esophagus and gastro-esophageal junction. *Ann Surg Oncol.* 2020;27:3182–3192.
42. Noble F, Lloyd MA, Turkington R, et al. Multicentre cohort study to define and validate pathological assessment of response to neoadjuvant therapy in oesophagogastric adenocarcinoma. *Br J Surg.* 2017;104:1816–1828.
43. Tu CC, Hsu PK, Chien LI, et al. Prognostic histological factors in patients with esophageal squamous cell carcinoma after preoperative chemoradiation followed by surgery. *BMC Cancer.* 2017;17:1–9.
44. Zeybek A, Erdogan A, Glkesen KH, et al. Significance of tumor length as prognostic factor for esophageal cancer. *Int Surg.* 2013;98:234–240.
45. Gu WS, Fang WZ, Liu CY, et al. Prognostic significance of combined pretreatment body mass index (BMI) and BMI loss in patients with esophageal cancer. *Cancer Manag Res.* 2019;11:3029–3041.
46. College of American Pathologists (CAP). Protocol for the Examination of Specimens From Patients With Carcinoma of the Esophagus.; 2017. Available at: [www.cap.org/cancerprotocols](http://www.cap.org/cancerprotocols). Accessed October 2, 2018.
47. Quinn LM, Hollis AC, Hodson J, et al. Prognostic significance of circumferential resection margin involvement in patients receiving potentially curative treatment for oesophageal cancer. *Eur J Surg Oncol.* 2018;44:1268–1277.
48. Low DE, Alderson D, Ceconello I, et al. International consensus on standardization of data collection for complications associated with esophagectomy: Esophagectomy Complications Consensus Group (ECCG). *Ann Surg.* 2015;262:286–294.
49. van der Werf LR, Busweiler LAD, van Sandick JW, et al. Reporting national outcomes after esophagectomy and gastrectomy according to the Esophageal Complications Consensus Group (ECCG). *Ann Surg.* 2020;271:1095–1101.