




# Systematic review and validation of clinical models predicting survival after oesophagectomy for adenocarcinoma

Piers R. Boshier<sup>1</sup> , Alison Swaray<sup>1</sup>, Bhamini Vadhvana<sup>1</sup>, Arun O'Sullivan<sup>1</sup>, Donald E. Low<sup>2</sup>, George B. Hanna<sup>1</sup> and Christopher J. Peters<sup>1\*</sup>, on behalf of POEM Biomarkers Consortium and OCCAMS Consortium

<sup>1</sup>Department of Surgery and Cancer, Imperial College London, London, UK

<sup>2</sup>Department of Thoracic Surgery, Virginia Mason Medical Centre, Seattle, Washington, USA

\*Correspondence to: Christopher J. Peters, Department of Surgery and Cancer, Imperial College London, 10th Floor, QEOM Wing, St Mary's Hospital, Praed Street, London W2 1NY, UK (e-mail: christopher.peters@imperial.ac.uk)

Presented as poster at the AUGIS Annual Meeting (Liverpool, 2019). Since this presentation the systematic literature search has been updated.

## Abstract

**Background:** Oesophageal adenocarcinoma poses a significant global health burden, yet the staging used to predict survival has limited ability to stratify patients by outcome. This study aimed to identify published clinical models that predict survival in oesophageal adenocarcinoma and to evaluate them using an independent international multicentre dataset.

**Methods:** A systematic literature search (title and abstract) using the Ovid Embase and MEDLINE databases (from 1947 to 11 July 2020) was performed. Inclusion criteria were studies that developed or validated a clinical prognostication model to predict either overall or disease-specific survival in patients with oesophageal adenocarcinoma undergoing surgical treatment with curative intent. Published models were validated using an independent dataset of 2450 patients who underwent oesophagectomy for oesophageal adenocarcinoma with curative intent.

**Results:** Seventeen articles were eligible for inclusion in the study. Eleven models were suitable for testing in the independent validation dataset and nine of these were able to stratify patients successfully into groups with significantly different survival outcomes. Area under the receiver operating characteristic curves for individual survival prediction models ranged from 0.658 to 0.705, suggesting poor-to-fair accuracy.

**Conclusion:** This study highlights the need to concentrate on robust methodologies and improved, independent, validation, to increase the likelihood of clinical adoption of survival predictions models.

## Introduction

Globally, oesophageal cancer affects 746 000 patients and is associated with 459 000 deaths annually<sup>1,2</sup>. While squamous cell carcinoma (SCC) remains the predominant global histological subtype, in recent years, the incidence of oesophageal adenocarcinoma has exceeded that of SCC in many parts of North America and Europe<sup>3</sup>.

Surgical resection of the oesophagus, with or without chemoradiotherapy or chemotherapy, offers the principal curative treatment modality for oesophageal adenocarcinoma but is only appropriate in a minority of patients who present with localized disease. For those patients who do undergo potentially curative surgical resection, overall 5-year survival is typically between 20 and 30 per cent and seldomly exceeds 50 per cent<sup>4,5</sup>. The associated morbidity and long-term sequelae of oesophageal resection serve as additional obstacles in the treatment of oesophageal cancer. The desire to establish greater equipoise of risk and benefit in the surgical management of oesophageal adenocarcinoma has meant that the ability to

accurately predict survival after oesophagectomy is of particular clinical significance.

The eighth edition of The American Joint Committee on Cancer (AJCC) and the International Union for Cancer Control (UICC) TNM classification, is principally based on anatomical tumour extent and remains the most widely adopted method of prognostication in oesophageal cancer<sup>6</sup>. Nevertheless, TNM staging criteria does not acknowledge other pathological, demographic, and clinical variables that are also known to impact upon survival<sup>7,8</sup>. In oesophageal cancer, as in other malignancies, predictive models of survival have been developed in an effort to improve prognostication. Such models are intended to support clinical decision-making and to better inform patients of their envisaged disease outcome. However, it is notable that few of these models are ever routinely used in clinical practice.

This systematic review aimed to identify published clinical and pathological models that predict survival in patients undergoing potentially curative surgical resection for oesophageal adenocarcinoma. Where possible, the performance of identified models were assessed using a prospectively collected multicentre dataset.

Received: October 14, 2021. Accepted: January 10, 2022

© The Author(s) 2022. Published by Oxford University Press on behalf of BJS Society Ltd.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

## Methods

### Search strategy

This systematic review was conducted in accordance with the recommendations of the Cochrane Library and MOOSE guidelines<sup>9</sup>. A systematic literature search using the Ovid Embase and MEDLINE databases (from 1947 to 11 July 2020) was performed to identify studies reporting predictive models of survival in oesophageal adenocarcinoma. Details of the search strategy are provided in [Table S1](#). The titles and abstracts of identified articles were screened by three independent reviewers (A.S., P.R.B., and B.V.) for potentially relevant studies that were subsequently subject to full-text review. To identify further potentially relevant studies, the reference lists of included articles were hand searched.

Inclusion criteria were studies that developed or validated a clinical prognostication model to predict either overall or disease-specific survival in patients with oesophageal adenocarcinoma undergoing surgical treatment with curative intent. Models that included patients with both adenocarcinoma and SCC were included if the former was the predominant tumour subtype used to develop the model. Likewise, studies including patients receiving therapies with curative intent, other than surgery, were included, if surgery was the predominant treatment modality within the study cohort. No restrictions were made regarding patient demographics, surgical approach, use of (neo)adjuvant therapies, or study design. Models developed for either pre- or postoperative use were also included. Models that included experimental metabolic and/or genetic biomarkers that are either not currently routinely available or used within clinical practice were excluded. Models that were developed through the use of artificial neural networks were also excluded, owing to their unsuitability for independent validation. Non-English-language articles and conference abstracts without an associated published full-text article were excluded. Any disagreement regarding a study's inclusion was resolved by a fourth reviewer (C.P.). Three

reviewers (A.S., B.V., and A.O'S.) independently extracted data from included studies.

### Definitions

Oesophageal adenocarcinoma was defined as a histologically specific malignancy affecting the oesophagus and/or gastroesophageal junction (Siewert type I and II) and oesophagectomy as surgery to resect all or part of the oesophagus through either an open, hybrid, or totally minimally invasive approach but not including endoscopic techniques. A prognostic model was defined as a multivariable tool designed to predict patient survival (overall or disease specific) or a surrogate factor that was shown by the authors to correlate directly and reliably with patient survival.

### Methodological quality assessment

The methodological quality of the included studies was assessed using the Critical Appraisal and Data Extraction for Systematic Reviews of Prediction Modelling Studies (CHARMS) checklist<sup>10</sup>. This checklist groups 35 key items into 11 domains that may be extracted from individual studies for the purpose of critical appraisal.

### Model validation

Models were validated using an independent dataset of 2450 patients who underwent oesophagectomy for oesophageal adenocarcinoma with curative intent. Data were acquired from the Oesophageal Cancer Clinical and Molecular Stratification Consortium (OCCAMS; 1088 patients)<sup>11,12</sup>, Predicting Outcomes of Esophageal Malignancy Biomarker Consortium (POEM; 811 patients), and from a high-volume North American Centre (Virginia Mason Medical Center; 551 patients). Local institutional review board approval was obtained by all participating centres for the purpose of sharing anonymised data. Characteristics of the validation dataset are provided in [Table 1](#).

**Table 1** Characteristics of validation dataset

	Validation cohort (n = 2450)		Available data/2450	
Mean (s.d.) age (years)	64.5 (10.5)		2442 (99.7)	
Sex (M:F)	2078:372 (84.8:15.2)		2450 (100.0)	
Mean (s.d.) BMI	27.5 (4.9)		841 (34.3)	
Median (i.q.r.) CCI	2.0 (2.0-3.0)		907 (37.0)	
ASA II/III/IV/V	482/478/8/1 (49.7/49.3/0.8/0.1)		969 (39.6)	
Adenocarcinoma	2450 (100.0)		2450 (100.0)	
Differentiation				
Well	127		2239 (91.4)	
Moderately	872			
Poor	1240			
Stage	Clinical	Pathological	Clinical	Pathological
Tis/0	15 (1.3)	104 (4.3)	1183 (48.3)	2431 (99.2)
T1	154 (13.0)	515 (21.2)		
T2	235 (19.9)	391 (16.1)		
T3	703 (61.7)	1330 (54.7)		
T4	49 (4.1)	91 (3.7)		
N0	455 (39.0)	979 (40.1)	1166 (47.6)	2441 (99.6)
N1	657 (56.3)	697 (28.6)		
N2	45 (3.9)	411 (16.8)		
N3	9 (0.8)	354 (14.5)		
Mean (s.d.) no. of lymph nodes excised	26.0 (18.2)	1967 (80.3)		
Mean (s.d.) no. of positive lymph nodes	3.4 (5.5)	1969 (80.4)		
V/N invasion	431 (46.5)	927 (37.8)		
R1	256 (20.0)	1279 (52.2)		
Neoadjuvant therapy	1061 (44.1)	2368 (96.7)		
Median (range) follow-up (months)	26 (0-245) <sup>a</sup>	2450 (100.0)		
Median survival (months)	39 (95%CI 36-43)	2450 (100.0)		

Data are n (%) unless otherwise stated. CCI, Charlson Comorbidity Index (non-age adjusted); V/N, vascular/neural invasion.

Table 2 Characteristics of included studies

Author	Year	Patient no. (training/validation)	Histology	Tumour site	Treatment	Outcome	Method of classification	References
Deans <i>et al.</i>	2007	220/—	AC/SCC	O/GOJ*	Surgery + any	OS	Nomogram	14
Lagarde <i>et al.</i>	2007	364/—	AC	O/GOJ	Surgery	DSS	Nomogram	15
Barbour <i>et al.</i>	2010	85/—	AC (T1)	O/GOJ	Surg	LNM, OS, DSS	Grades (I–IV)	16
Langer <i>et al.</i>	2014	360/—	AC	O	Surgery + nC†	OS	Grades (A, B, C)	17
Eil <i>et al.</i>	2014	824/int‡	AC/SCC	O	Surgery ± nCRT	OS	Risk calculator	18
Shapiro <i>et al.</i>	2016	626/int‡	AC/SCC	O/GOJ	Surgery + nCRT	OS	Nomogram	19
Davison <i>et al.</i>	2016	210/39	AC (T1)	O/GOJ	Surgery	LNM, OS, TRG	Grades (I–IV)	20
Cao <i>et al.</i>	2016	4109/145	AC/SCC	O	Surgery ± nRT	OS	Nomogram	21
Lindenmann <i>et al.</i>	2017	174/—	AC/SCC	O	Surgery + nCRT	OS, PFS	Grade (I, II)	22
Zhou <i>et al.</i>	2015	953/181	AC	GOJ	Surgery+?	OS	Nomogram	23
Gabriel <i>et al.</i>	2017	7179/1795	AC	O	Surgery ± nCRT	OS	Risk calculator	24
Zhang <i>et al.</i>	2017	355/—	AC	GOJ	Surgery	OS	Nomogram	25
Xie <i>et al.</i>	2018	1948/476	AC/SCC	O/GOJ	Surgery + nRT	DSS	Nomogram	26
Goense <i>et al.</i>	2018	373/—	AC	O	Surgery + nCRT	PFS, OS	Nomogram	27
Liu <i>et al.</i>	2019	1090/728	AC	GOJ	Surgery + nRT	OS	Nomogram	28
Hagens <i>et al.</i>	2020	660/—	AC/SCC	O	Surgery + nCRT	OS	Nomogram	29
Du <i>et al.</i>	2020	3198/1368	AC/SCC	O	Surgery ± nCRT	DSS, OS	Nomogram	30

\*Includes gastric cancer. †Thirty-one patients also received neoadjuvant chemotherapy. ‡Int, internal validation; nC, neoadjuvant chemotherapy; nCRT, neoadjuvant chemoradiotherapy; nRT, neo-adjuvant radiotherapy; AC, adenocarcinoma; SCC, squamous cell carcinoma; O, oesophageal; GOJ, gastroesophageal junction; OS, overall survival; DSS, disease-specific survival; LNM, lymph node metastasis; TRG, tumour regression grade; PFS, progression-free survival.

Table 3 Comparison of survival outcomes

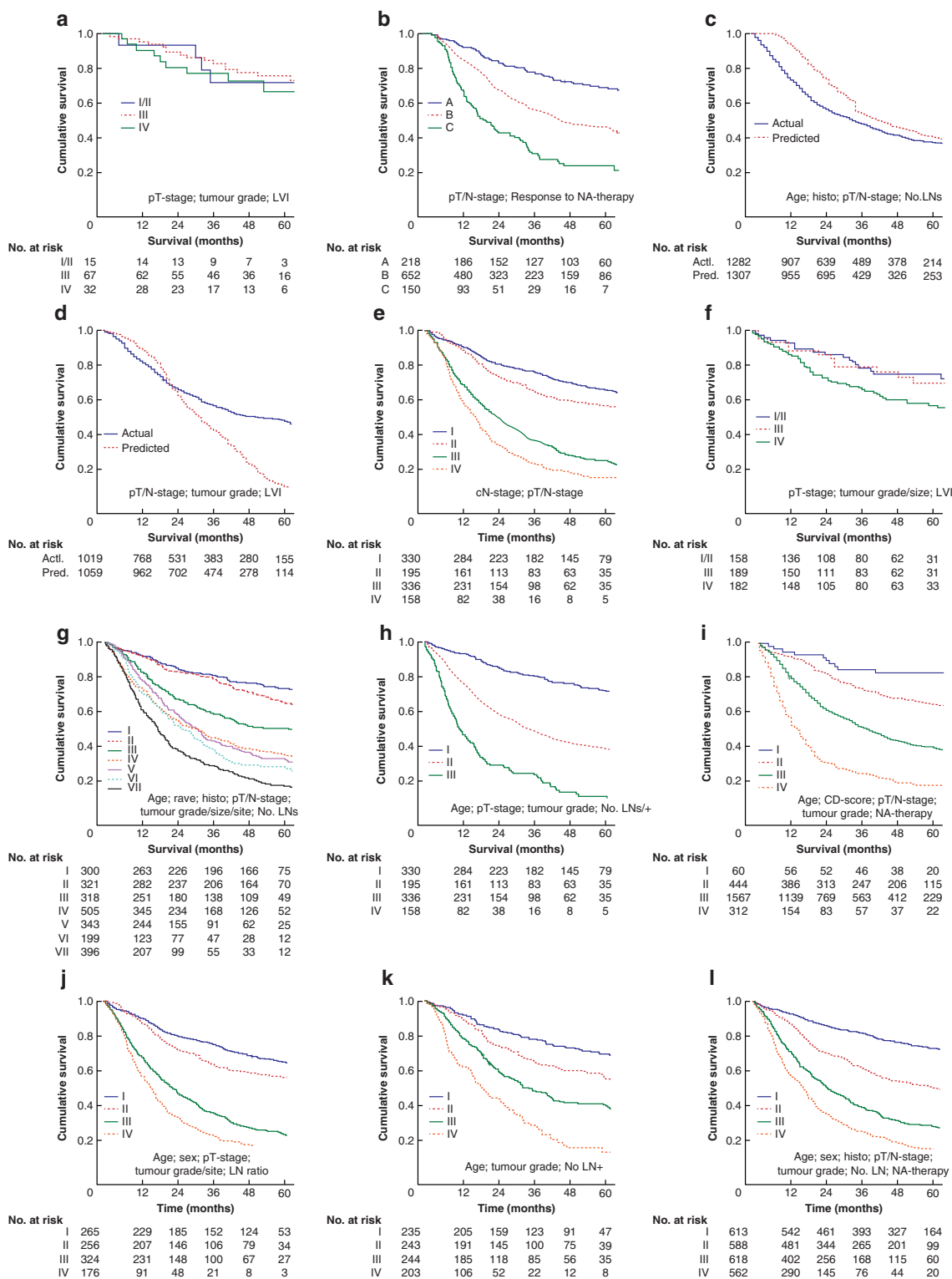
	Original cohort			Validation cohort			P	AUC†	
	No.	OS (%)*	5-year OS for model subgroups (%)	No.	Median survival (months)†	5-year OS (%)			5-year OS for model subgroups (%)
Barbour <i>et al.</i>	85	77	I 97 II 87 III 50 IV 29	114	137 (121–153)	75	I/II 73 — III 78 IV 72	0.859	—
Langer <i>et al.</i>	360	—	A 64 B 42 C 18	1061	48 (40–57)	57	A 73 B 57 C 33	<0.001	—
Eil <i>et al.</i>	824	—	—	2450	39 (36–43)	49	—	—	—
Shapiro <i>et al.</i>	626	—	—	1061	48 (40–57)	57	—	<0.001	0.672 (0.639–0.705) P < 0.001
Davison <i>et al.</i>	210	—	I 89 II 87 III 65 IV 42	544	102 (84–120)	69	I/II 77 — II 74 III 63	<0.001	—
Cao <i>et al.</i>	4109	40	I 84 II 69 III 54 IV 40 V 22 VI 11 VII 5	2450	39 (36–43)	49	I 75 II 69 III 54 IV 40 V 42 VI 39 VII 33	<0.001	0.658 (0.637–0.680) P < 0.001
Zhou <i>et al.</i>	953	38	I 65 II 37 III 18	2450	39 (36–43)	49	I 76 II 46 III 27	<0.001	0.696 (0.675–0.717) P < 0.001
Gabriel <i>et al.</i>	7179	54‡	—	2450	39 (36–43)	56	—	<0.001	0.68 (0.661–0.703) P < 0.001
Xie <i>et al.</i>	1948	—	—	1061	48 (40–57)	57	—	<0.001	0.673 (0.640–0.705) P < 0.001
Liu <i>et al.</i>	1090	—	—	964	47 (38–56)	57	—	<0.001	0.677 (0.643–0.711) P < 0.001
Du <i>et al.</i>	3198	41	—	2450	39 (36–43)	49	—	<0.001	0.705 (0.684–0.726) P < 0.001

\*Values are for 5-year overall survival (OS) unless otherwise stated. †Values in parentheses are 95% confidence intervals. ‡Value is for survival at 3 years. AUC, area under the receiver operating characteristic curve.

## Statistical analysis

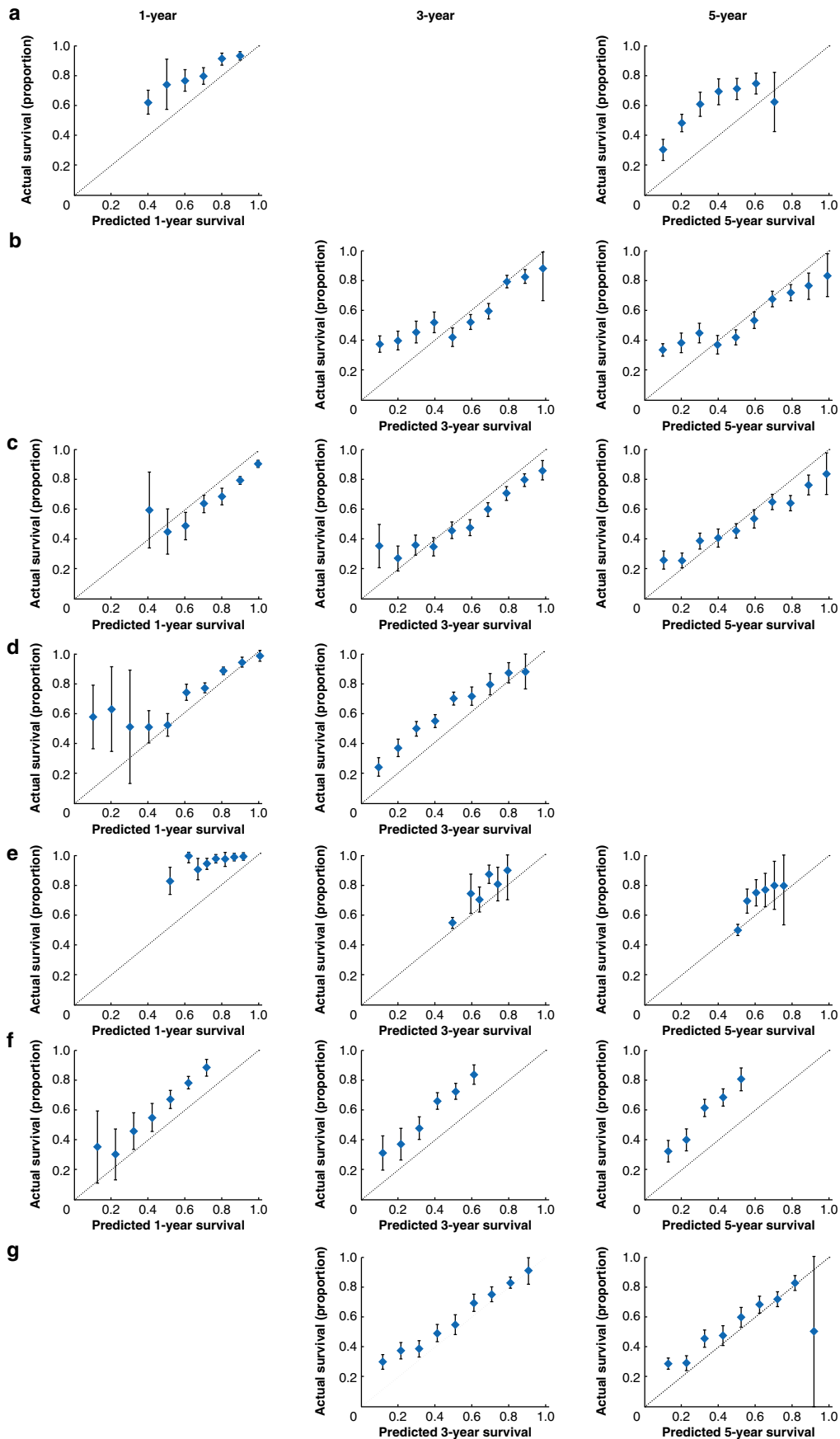
Model validation was dependent on the concordance of test variables between the model and available variables within the validation dataset. Validation was performed according to the

published eligibility criteria within each study. Missing data were dealt with by imputing the mode for categorical variables, and the median for continuous variables<sup>13</sup>. Risk stratification scoring systems were assessed using Kaplan–Meier curves and



**Fig. 1** Kaplan-Meier plots of predicted survival

**a** Barbour *et al.*, stratification of survival based on nodal metastasis risk grade (grades I/II were considered as a single grade;  $P = 0.859$ )<sup>16</sup>. **b** Langer *et al.*, stratification of survival based on prognosis score ( $A = 3, B = 4-5, C = 6; P < 0.001$ )<sup>17</sup>. **c** Eil *et al.*, predicted and actual survival of patients who underwent surgery alone ( $P < 0.001$ )<sup>18</sup>. **d** Eil *et al.*, predicted and actual survival of patients who underwent neoadjuvant therapy and surgery ( $P < 0.001$ )<sup>18</sup>. **e** Shapiro *et al.*, stratification of survival based on nomogram score ( $I = 0-4, II = 5-8, III = 9-11, IV \geq 12; P < 0.001$ )<sup>19</sup>. **f** Davison *et al.*, stratification of survival based on nodal metastasis risk grade (grades I/II were considered as a single grade;  $P < 0.009$ )<sup>20</sup>. **g** Cao *et al.*, stratification of survival based on nomogram score ( $I = 0.0-6.6, II = 6.7-9.0, III = 9.1-11.0, IV = 11.1-13.8, V = 13.9-16.7, VI = 16.8-18.7, VII \geq 18.8; P < 0.001$ )<sup>21</sup>. **h** Zhou *et al.*, stratification of survival based on nomogram score ( $I = 0-93, II = 94-187, III = 188-280; P < 0.001$ )<sup>23</sup>. **i** Gabriel *et al.*, stratification of survival based on nomogram score ( $I = 0-88, II = 89-142, III = 143-172, IV \geq 173; P < 0.001$ )<sup>26</sup>. **k** Liu *et al.*, stratification of survival based on nomogram score ( $I = 0-48, II = 49-64, III = 65-88, IV \geq 89; P < 0.001$ )<sup>28</sup>. **l** Du *et al.*, stratification of survival based on nomogram score ( $I = 0-99, II = 100-158, III = 159-200, IV \geq 200; P < 0.001$ )<sup>30</sup>. Box contains variables included in the model. pT-stage, pathological T-stage. pN-stage, pathological N-stage. LVI, lymphovascular invasion. NA-therapy, neoadjuvant therapy. Histo, histology. No. LNs, number of lymph nodes examined. No LNs/+, total number of lymph nodes examined and number of positive lymph nodes. CD-score, Charlson-Deyo comorbidity score



**Fig. 2** Calibration plots curves for predicted survival

a Shapiro et al.<sup>19</sup>, b Cao et al.<sup>21</sup>, c Zhou et al.<sup>23</sup>, d Gabriel et al.<sup>24</sup>, e Xie et al.<sup>26</sup>, f Liu et al.<sup>28</sup>, g Du et al.<sup>30</sup>

the log-rank test. Individualized prediction scoring models were evaluated by plotting calibration curves of predicted against actual survival. Where possible, model performance was further evaluated using receiver operating characteristic (ROC) curves and the corresponding area under the curve (AUC). Statistical analysis was performed using SPSS statistics version 27.0 (IBM, Armonk, New York, USA), with  $P < 0.05$  considered to signify statistical significance.

## Results

Of the 8133 studies identified through the electronic search 17 were eligible for inclusion<sup>14–30</sup>. A PRISMA diagram of study selection is provided in Fig. S1. Included studies cumulatively assessed 27460 patients during the period 1988 to 2019. The characteristics of the included studies are presented in Table 2. Further study characteristics and details of variables assessed and utilized within models are provided in Tables S2 and S3. The commonest features included in models to predict survival were T stage (11 studies), tumour grade (10 studies), N stage (nine studies), and patient age (eight studies). In five studies, patient cohorts were prospectively sourced solely for the development of the intended prognostic models. The remaining studies used pre-existing data registries, including the Surveillance, Epidemiology, and End Results Program (SEER)<sup>18,21,23,26,28,30</sup>, the National Cancer Database (NCDB)<sup>24</sup>, trial datasets<sup>19</sup>, and internal hospital datasets<sup>22,31</sup>. The inclusion of two studies in which the proposed model was intended to predict lymph node metastasis<sup>16,20</sup> was based on the concurrent association of this outcome with survival that was demonstrated by the authors.

## Appraisal of studies

Critical appraisal of included studies was performed in accordance with the CHARMS checklist. Eleven studies adopted either a prospective or retrospective cohort design, while seven studies used data from a national data registry. One study used a retrospective case-control design. A summary of the risk of bias within included studies, as determined by the CHARMS checklist, is provided in Table S4. Notably, none of the 17 papers achieved the highest CHARMS standards in terms of sample size, missing data, and model development, and only one paper achieved the best rating in candidate predictors. In contrast, interpretation and discussion were performed well in 15 of 17 studies.

## Assessment of model performance

Of the 17 studies identified from the literature search, 11 presented models that were suitable for testing against the independent validation dataset<sup>16–21,23,24,26,28,30</sup>. Details of models that were assessed, including a comparison of survival outcomes with the validation dataset, are presented in Table 3 and Figs 1 and 2, with full details provided in Table S5. With the exception of the models reported by Barbour et al.<sup>16</sup> (Fig. 1a) and Eil et al.<sup>18</sup> (Fig. 1c,d), all models were able to predict survival successfully in patients within the independent validation dataset. The model published by Davison et al.<sup>20</sup> demonstrated poor discrimination of groups I/II and III versus group IV (Fig. 1f). In terms of individual predictions of survival, this was, generally speaking, more accurate for longer surviving patients, with ROC-AUC ranging from 0.658 to 0.705 (Table 3); however, some models tended to under-predict survival<sup>19,24</sup>.

## Discussion

This is the first study to systematically identify and attempt to validate published models for the prediction of survival in oesophageal adenocarcinoma. Owing to the generally poor prognosis of oesophageal cancer<sup>2</sup>, it is vital to have accurate prognostic information so that patients and clinicians can make informed treatment choices. This makes a clear case for the benefit of well-constructed and reliable predictive models. While the TNM system stratifies patients into groups with significantly different outcomes<sup>32</sup>, the majority of patients fall into stage III, limiting its real-world utility. This systematic review identified 17 published prediction models, that assessed a combined 50 variables. These studies were of variable methodological quality, with none reaching the highest standards in several core assessment criteria: sample size, missing data, and model development<sup>10</sup>. Ten of the 17 models were derived from existing cohorts of patients and seven from national registries.

Eleven of the 17 models could be tested using a large multicentre cohort of patients from the UK, USA, Ireland, and the Netherlands, including cohorts from the OCCAMS Consortium<sup>11,12</sup> and the newly formed POEM Biomarkers Consortium. This mixed cohort represents a good test of their predictive power. It was reassuring that nine of the 11 models successfully validated; however, complete separation of all prognostic groups was not seen in two studies<sup>16,20</sup>. It is also notable that one of the models that failed to validate<sup>16</sup> had one of the smallest assessment cohorts (85 patients) and was one of the oldest studies (1991 to 2008).

Models that predicted individual survival did reasonably well but with a tendency to under-predict short-term survival. The AUCs for the models clustered around 0.65 to 0.70, which is considered poor-to-fair accuracy. Therefore, this raises the question of how these models could be used in clinical scenarios. It is noteworthy that none of the models presented herein is currently in routine clinical use. Existing models may have limited use on an individual patient basis but could potentially be used to stratify patients into different treatment groups: neoadjuvant chemotherapy and surgery versus surgery alone, for example. It is notable that many were designed for a subset of patients with oesophageal adenocarcinoma (e.g. T1 or neoadjuvant chemoradiotherapy) that potentially makes them more restricted in their utility. However, in one such model, Davison et al. were able to subdivide the T1 group to some degree, identifying a poorer prognosis group (IV) that may warrant more aggressive treatment<sup>20</sup>.

A further limitation of existing models is, for the most part, their reliance upon knowledge of pathological staging. For patients and clinicians, it would be more informative to have information regarding prognosis at the time of diagnosis, when it may have the largest impact on clinical decision-making and treatment planning. Future work should therefore place a greater emphasis on the identification of pretreatment prognostic markers.

The strengths of this study include the carefully constructed systematic review following published Cochrane Library and MOOSE guidelines<sup>9</sup>. The validation cohort was large, mixed, and multinational; however, as the individual databases differed in terms of the data recorded it was not possible to use all the patients to validate every model. In particular, the small numbers of patients used to test the model published by Barbour et al.<sup>16</sup> may have been linked to its failure to validate; however, the test set was larger than the cohort used to generate the original model. It was not possible to assess the performance of all identified models, owing to a limited number of variables within the validation dataset. A further

limitation was the decision to exclude models that included novel biomarkers such as immunohistochemistry and genetic markers<sup>11</sup>. Such models are likely to have an important role to play in the future, but it was decided to concentrate on those models that could be immediately implemented by surgical centres and therefore, by definition, were models that could be tested easily.

This systematic review and validation of models designed to predict survival in oesophageal adenocarcinoma demonstrates that the models already have potential to stratify patients in a more granular way than standard TNM staging. However, none has yet achieved widespread adoption, with this study being the first time any of these models have been tested in other cohorts. To develop a robust and reliable model it is important to avoid as many of the potential sources of bias as possible and to generate and test the model in a large and multicentre cohort. This avoids the danger of overfitting the data to local outcomes<sup>33</sup>. While these 17 models have added to the field, it has not been translated into widespread adoption, and therefore they have failed to alter management decisions or improve outcomes in the real world. This must be the goal for any predictive model. While we did not include models that included biomarkers it is notable that, despite many being proposed<sup>34</sup>, none of these are in widespread use either, very likely for the same reasons that these clinical models have not been adopted.

Future work developing and validating predictive models in oesophageal cancer must concentrate on adopting robust and bias-free methodologies, suitably sized and representative patient cohorts, and, most importantly, externally validate the models in independent patient groups. If the work is carried out in this way, it will increase the likelihood of adoption and therefore improve the chance of improving patients' outcomes.

## Collaborators

We are very grateful to the POEMS Biomarker and OCCAMS Consortium collaborators that provided anonymised data for this study and reviewed the final manuscript: R.C. Fitzgerald, R.H. Hardwick, J.M.H. Bennett, P.M. Safranek, A.C. Hindmarsh, N. Carroll, and J.R. O'Neill (Cambridge Oesophagogastric Centre/Cambridge University); T. Underwood, R. Walker, and J. Harrington (Southampton University); R. Turkington (Queens University Belfast); K.S. Nason, J.M. Davison, and J.D. Luketich (University of Pittsburgh); J. Reynolds and J. O'Sullivan (Trinity St. James's Cancer Institute, Dublin); R. Skipworth (Edinburgh Royal Infirmary); S. Darnton and R. Steyn (Birmingham Heartlands Hospital); J. Going and M. McKernan (Glasgow Royal Infirmary); R. Stuart (Ross Hall Hospital, Glasgow); M. Moorghen, J. Blazeby, and C. P. (Barham Bristol Royal Infirmary); J. Shapiro, W.N.M. Dinjens, K. Biermann, and B.P.L. Wijnhoven (Erasmus University Medical Center, Rotterdam); and C. Rajaguru, N. Imrit, and N. Maynard (Oxford Radcliffe Hospitals NHS Trust).

## Funding

G.B.H. and C.J.P. are supported by the NIHR Imperial Biomedical Research Centre, which also provided infrastructure support. P.R.B. is an NIHR clinical lecturer. OCCAMS2 was funded by a Programme Grant from Cancer Research UK (RG81771/84119).

## Acknowledgements

The Predicting Outcome in Esophageal Malignancy (POEM) Biomarkers Consortium is an international network of centres

aiming to improve the development and validation of biomarkers in oesophageal cancer. The Oesophageal Cancer Clinical and Molecular Stratification (OCCAMS) Consortium is a network of UK clinical centres recruiting oesophageal cancer patients for tissue and data collection. This information is used to identify clinical, demographic, and molecular factors affecting disease progression.

**Disclosure.** The authors declare no conflict of interest.

## Supplementary material

Supplementary material is available at BJS online.

## References

1. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990–2016: a systematic analysis for the global burden of disease study 2016. *Lancet* 2017;**390**:1211–1259.
2. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet* 2017;**390**:1151–1210.
3. Hur C, Miller M, Kong CY, Dowling EC, Nattinger KJ, Dunn M *et al*. Trends in esophageal adenocarcinoma incidence and mortality. *Cancer* 2013;**119**:1149–1158.
4. National Cancer Institute. Surveillance, Epidemiology, and End Results Program. Cancer Stat Facts: Esophageal Cancer, <https://seer.cancer.gov/statfacts/html/esoph.html> (accessed 21 September 2017).
5. Shapiro J, van Lanschot JJB, Hulshof MCCM, van Hagen P, van Berge Henegouwen MI, Wijnhoven BPL *et al*. Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. *Lancet Oncol* 2015;**16**:1090–1098.
6. Rice T, Kelsen DP, Blackstone EH, Ishwaran H, Patil DT, Bass AJ *et al*. Esophagus and esophagogastric junction. In: Amin MA, Edge S, Greene FL, Schilsky RL, Gaspar LE (eds), *AJCC Cancer Staging Manual*, 8th ed. New York: Springer Science+Business Media, 2016, 185–202.
7. Crumley AB, McMillan DC, McKernan M, Going JJ, Shearer CJ, Stuart RC. An elevated C-reactive protein concentration, prior to surgery, predicts poor cancer-specific survival in patients undergoing resection for gastro-oesophageal cancer. *Br J Cancer* 2006;**94**:1568–1571.
8. Mariette C, Piessen G, Briez N, Triboulet JP. The number of metastatic lymph nodes and the ratio between metastatic and examined lymph nodes are independent prognostic factors in esophageal cancer regardless of neoadjuvant chemoradiation or lymphadenectomy extent. *Ann Surg* 2008;**247**:365–371.
9. Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D *et al*. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis of observational studies in epidemiology (MOOSE) group. *JAMA* 2000;**283**:2008–2012.
10. Moons KG, de Groot JAH, Bouwmeester W, Vergouwe Y, Mallett S, Altman DG *et al*. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med* 2014;**11**:e1001744.
11. Peters CJ, Rees JR, Hardwick RH, Hardwick JS, Vowler SL, Ong CA *et al*. A 4-gene signature predicts survival of patients with

- resected adenocarcinoma of the esophagus, junction, and gastric cardia. *Gastroenterology* 2010;**139**:1995–2004.
12. Ong CA, Shapiro J, Nason KS, Davison JM, Liu X, Ross-Innes C et al. Three-gene immunohistochemical panel adds to clinical staging algorithms to predict prognosis for patients with esophageal adenocarcinoma. *J Clin Oncol* 2013;**31**:1576–1582.
  13. 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**:361–387.
  14. Deans DA, Wigmore SJ, de Beaux AC, Paterson-Brown S, Garden OJ, Fearon KC. Clinical prognostic scoring system to aid decision-making in gastro-oesophageal cancer. *Br J Surg* 2007;**94**:1501–1508.
  15. Lagarde SM, Reitsma JB, de Castro SMM, Ten Kate FJW, Busch ORC, van Lanschot JJB. Prognostic nomogram for patients undergoing oesophagectomy for adenocarcinoma of the oesophagus or gastro-oesophageal junction. *Br J Surg* 2007;**94**:1361–1368.
  16. Barbour AP, Jones M, Brown I, Gotley DC, Martin I, Thomas J et al. Risk stratification for early esophageal adenocarcinoma: analysis of lymphatic spread and prognostic factors. *Ann Surg Oncol* 2010;**17**:2494–2502.
  17. Langer R, Becker K, Zlobec I, Gertler R, Sisic L, Buchler M et al. A multifactorial histopathologic score for the prediction of prognosis of resected esophageal adenocarcinomas after neoadjuvant chemotherapy. *Ann Surg Oncol* 2014;**21**:915–921.
  18. Eil R, Diggs BS, Wang SJ, Dolan JP, Hunter JG, Thomas CR. Nomogram for predicting the benefit of neoadjuvant chemoradiotherapy for patients with esophageal cancer: a SEER-Medicare analysis. *Cancer* 2014;**120**:492–498.
  19. Shapiro J, van Klaveren D, Lagarde SM, Toxopeus ELA, van der Gaast A, Hulshof MCCM et al. Prediction of survival in patients with oesophageal or junctional cancer receiving neoadjuvant chemoradiotherapy and surgery. *Br J Surg* 2016;**103**:1039–1047.
  20. Davison JM, Landau MS, Luketich JD, McGrath KM, Foxwell TJ, Landsittel DP et al. A model based on pathologic features of superficial esophageal adenocarcinoma complements clinical node staging in determining risk of metastasis to lymph nodes. *Clin Gastroenterol Hepatol* 2016;**14**:369–377.
  21. Cao J, Yuan P, Wang L, Wang Y, Ma H, Yuan X et al. Clinical nomogram for predicting survival of esophageal cancer patients after esophagectomy. *Sci Rep* 2016;**6**:26684.
  22. Lindenmann J, Fink-Neuboeck N, Avian A, Pichler M, Habitzruther M, Maier A et al. Preoperative Glasgow prognostic score as additional independent prognostic parameter for patients with esophageal cancer after curative esophagectomy. *Eur J Surg Oncol* 2017;**43**:445–453.
  23. Zhou Z, Zhang H, Xu Z, Li W, Dang C, Song Y. Nomogram predicted survival of patients with adenocarcinoma of esophago-gastric junction. *World J Surg Oncol* 2015;**13**:197.
  24. Gabriel E, Attwood K, Shah R, Nurkin S, Hochwald S, Kukar M. Novel calculator to estimate overall survival benefit from neoadjuvant chemoradiation in patients with esophageal adenocarcinoma. *J Am Coll Surg* 2017;**224**:884–894.
  25. Zhang L, Su Y, Chen Z, Wei Z, Han W, Xu A. The prognostic value of preoperative inflammation-based prognostic scores and nutritional status for overall survival in resected patients with nonmetastatic Siewert type II/III adenocarcinoma of esophago-gastric junction. *Medicine* 2017;**96**:e7647.
  26. Xie K, Liu S, Liu J. Nomogram predicts survival benefit for non-metastatic esophageal cancer patients who underwent preoperative radiotherapy. *Cancer Manag Res* 2018;**10**:3657–3668.
  27. Goense L, van Rossum PSN, Xi M, Maru DM, Carter BW, Meijer GJ et al. Preoperative nomogram to risk stratify patients for the benefit of trimodality therapy in esophageal adenocarcinoma. *Ann Surg Oncol* 2018;**25**:1598–1607.
  28. Liu F, Zhou R, Jiang F, Liu G, Li K, Zhu G. Proposal of a nomogram for predicting survival in patients with siewert type ii adenocarcinoma of the esophagogastric junction after preoperative radiation. *Ann Surg Oncol* 2019;**26**:1292–1300.
  29. Hagens ERC, Feenstra ML, Eshuis WJ, Hulshof MCCM, van Laarhoven HWM, van Berge Henegouwen MI et al. Conditional survival after neoadjuvant chemoradiotherapy and surgery for oesophageal cancer. *Br J Surg* 2020;**107**:1053–1061.
  30. Du F, Sun Z, Jia J, Yang Y, Yu J, Shi Y et al. Development and validation of an individualized nomogram for predicting survival in patients with esophageal carcinoma after resection. *J Cancer* 2020;**11**:4023–4029.
  31. Gockel I, Niebisch S, Campbell LK, Sgourakis G, Junginger T. Prognostic scoring system predictive of survival after surgical resection of esophageal carcinoma. *Thorac Cardiovasc Surg* 2013;**61**:470–478.
  32. Rice TW, Rusch VW, Ishwaran H, Blackstone EH, Worldwide Esophageal Cancer Collaboration. 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 Cancer Staging Manuals. *Cancer* 2010;**116**:3763–3773.
  33. Cook JA, Ranstam J. Overfitting. *Br J Surg* 2016;**103**:1814.
  34. Matthews LHM, Noble F, Tod J, Jaynes E, Harris S, Primrose JN et al. Systematic review and meta-analysis of immunohistochemical prognostic biomarkers in resected oesophageal adenocarcinoma. *Br J Cancer* 2015;**113**:107–118.