



University of Dundee

Multicentre validation of CT grey-level co-occurrence matrix features for overall survival in primary oesophageal adenocarcinoma

O'Shea, Robert; Withey, Samuel J.; Owczarczyk, Kasia; Rookyard, Christopher; Gossage, James; Godfrey, Edmund

Published in:
European Radiology

DOI:
[10.1007/s00330-024-10666-y](https://doi.org/10.1007/s00330-024-10666-y)

Publication date:
2024

Licence:
CC BY

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

O'Shea, R., Withey, S. J., Owczarczyk, K., Rookyard, C., Gossage, J., Godfrey, E., Jobling, C., Parsons, S. L., Skipworth, R. J. E., Goh, V., OCCAMS Consortium, & Petty, R. D. (2024). Multicentre validation of CT grey-level co-occurrence matrix features for overall survival in primary oesophageal adenocarcinoma. *European Radiology*, 34(10), 6919-6928. <https://doi.org/10.1007/s00330-024-10666-y>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.


Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

GASTROINTESTINAL



Multicentre validation of CT grey-level co-occurrence matrix features for overall survival in primary oesophageal adenocarcinoma

Robert O'Shea¹, Samuel J. Withey^{1,2}, Kasia Owczarczyk^{1,3}, Christopher Rookyard¹, James Gossage⁴, Edmund Godfrey⁵, Craig Jobling⁶, Simon L. Parsons⁷, Richard J. E. Skipworth⁸, Vicky Goh^{1,9*}  on behalf of the OCCAMS Consortium

Abstract

Background Personalising management of primary oesophageal adenocarcinoma requires better risk stratification. Lack of independent validation of proposed imaging biomarkers has hampered clinical translation. We aimed to prospectively validate previously identified prognostic grey-level co-occurrence matrix (GLCM) CT features for 3-year overall survival.

Methods Following ethical approval, clinical and contrast-enhanced CT data were acquired from participants from five institutions. Data from three institutions were used for training and two for testing. Survival classifiers were modelled on prespecified variables ('Clinical' model: age, clinical T-stage, clinical N-stage; 'ClinVol' model: clinical features + CT tumour volume; 'ClinRad' model: ClinVol features + GLCM_Correlation and GLCM_Contrast). To reflect current clinical practice, baseline stage was also modelled as a univariate predictor ('Stage'). Discrimination was assessed by area under the receiver operating curve (AUC) analysis; calibration by Brier scores; and clinical relevance by thresholding risk scores to achieve 90% sensitivity for 3-year mortality.

Results A total of 162 participants were included (144 male; median 67 years [IQR 59, 72]; training, 95 participants; testing, 67 participants). Median survival was 998 days [IQR 486, 1594]. The ClinRad model yielded the greatest test discrimination (AUC, 0.68 [95% CI 0.54, 0.81]) that outperformed Stage (Δ AUC, 0.12 [95% CI 0.01, 0.23]; $p = .04$). The Clinical and ClinVol models yielded comparable test discrimination (AUC, 0.66 [95% CI 0.51, 0.80] vs. 0.65 [95% CI 0.50, 0.79]; $p > .05$). Test sensitivity of 90% was achieved by ClinRad and Stage models only.

Conclusions Compared to Stage, multivariable models of prespecified clinical and radiomic variables yielded improved prediction of 3-year overall survival.

Clinical relevance statement Previously identified radiomic features are prognostic but may not substantially improve risk stratification on their own.

Key Points

- Better risk stratification is needed in primary oesophageal cancer to personalise management.

*Correspondence:

Vicky Goh

vicky.goh@kcl.ac.uk

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

- *Previously identified CT features—GLCM_Correlation and GLCM_Contrast—contain incremental prognostic information to age and clinical stage.*
- *Compared to staging, multivariable clinoradiomic models improve discrimination of 3-year overall survival.*

Keywords Oesophageal neoplasms, Adenocarcinoma, Radiomics, Prognosis, Precision medicine

Introduction

Oesophageal cancer presents a major burden worldwide [1]. A multimodal treatment approach (surgery with chemotherapy or chemoradiotherapy) is standard of care following landmark trials [2–4] and offers the best chance of survival for resectable cancer. However, despite this, outcome remains poor for patients treated with curative intent with mortality rates of 45–53% in the first year post diagnosis [5] and median post-progression survival of only 13 months [6]. There is a greater need for personalisation of management to obviate treatment in patients who may not benefit substantially. Alongside this, with the growing interest in total neoadjuvant therapy, and ongoing trials of perioperative immunotherapy, better initial risk stratification of patients at diagnosis is needed to guide management.

Currently, the clinical TNM (tumour-node-metastasis) stage guides management [7] but it has a low predictive accuracy with contemporaneous pathological stage for early stage cancers [8], as well as limitations in prognostication. Prognostic information has central importance for patient decision-making, with cancer patients ranking prognosis, diagnosis and treatment options as their highest information priority [9].

Modelling studies have demonstrated scope to improve upon prognostication in oesophageal adenocarcinoma, including through CT imaging radiomic approaches [10–12]. Initial publications have highlighted the potential of different locoregional features. For example, Piazzese et al identified an association of CT grey-level zone distance variance with overall survival in a multicentre cohort ($n = 213$), which was independent of dimensionality and contrast administration [11], whilst Larue et al developed a random forest radiomic model including 40 CT features to predict 3-year overall survival (OS) in oesophageal cancer patients treated with chemoradiotherapy ($n = 239$) [12].

To date, clinical translation of imaging radiomic models has been hampered by a relative paucity of independent external validation. Thus, we aimed to validate a prognostic model for 3-year OS including pre-specified clinical features and previously proposed CT radiomic features in a prospective multicentre setting for patients with primary oesophageal adenocarcinoma planned for curative treatment.

Methods

Participants and datasets

Following ethical approval, clinical data and CT imaging were obtained prospectively from five institutions participating in the OCCAMS (Oesophageal Cancer Clinical and Molecular Stratification) Consortium. Consecutive participants with non-metastatic, pathologically proven oesophageal adenocarcinoma who underwent staging contrast-enhanced CT imaging and planned for definitive treatment were eligible. Participants were excluded if (1) no tumour was visible on CT; (2) CT images were unavailable/corrupted; and (3) concurrent malignancy was present. Data from three institutions were used for model development and two institutions for model testing.

CT imaging and analysis

Contrast-enhanced CT was performed according to institutional practice and included arterial phase imaging of the thorax and upper abdomen. CT acquisition and reconstruction parameters for the training and test datasets are summarised in Supplementary Table 1. Pre-processing of CT images was undertaken as per Image Biomarker Standardisation Initiative recommendations [13]. CT slice thickness was linearly interpolated to 2 mm and attenuation values were converted to Hounsfield units with PyDICOM [14]. The primary tumour was segmented by a radiologist (with 5 years' experience), who was blinded to clinical outcomes, on the arterial phase thoraco-abdominal CT images. The rationale for using the arterial versus portal venous phase was prior studies showing better tumour conspicuity [15] and tumour staging accuracy [16] for the arterial phase. Tumour segmentations were reviewed and adjusted as required by a second radiologist (with > 20 years' experience). Examples of tumour segmentations are shown in Supplementary Figs. 1 and 2. 3D Radiomic features were extracted from the segmented volume of interest using PyRadiomics version 3.0.1 under default parameters (no image filters, no normalisation, no voxel array shift, grey-level discretisation at fixed bin width of 25 Hounsfield units) [17].

Radiomic feature selection

Following recommendations [18], radiomic features were pre-selected based on previous published studies. This

obviated data-driven feature selection thereby reducing the risk of data overfitting. Published studies were evaluated using the 'Transparent reporting of a multivariable prediction model for individual prognosis' (TRIPOD) [19] and 'Radiomics Quality Score' checklists [20]. Identifiable Image Biomarker Standardisation Initiative features were sought [13].

Of the identified published studies [11, 12, 21, 22], features proposed by Piazzese et al were excluded as the interpolation strategy employed (2 mm isotropic) would have resulted in a significant reduction of the axial image resolution [11]. Features proposed by Larue et al were excluded as they evaluated nonlinear effects in a high-dimensional feature set using a random forest method, complicating the extraction of a few individually informative linear predictions [12]. Zhang et al achieved a high TRIPOD score of 30 [21] and examined a limited number of previously proposed predictors, each of which was identifiable; thus, these features were extracted—GLCM_Contrast, GLCM_Correlation and GLCM_InverseDifferenceMoment. Furthermore, GLCM correlation was independently identified as a predictor of response in Klaasen et al [22]. Following Peduzzi and Concato's guideline recommendation of more than 10 events per modelled feature [23], an unsupervised method was applied to identify the most suitable two of the three proposed features. The most collinear GLCM feature with respect to tumour volume and the other GLCM features was excluded.

Statistical analysis

Differences of participant characteristics between training and testing datasets were tested with the Fisher test for categorical variables or the two-sided *t*-test for continuous variables. For modelling, unpenalised logistic regression models were fitted to predict 3-year OS using base R. With a median survival of 24 months following surgery alone and 46 months following neoadjuvant chemoradiotherapy and surgery, 3-year OS is a meaningful endpoint in oesophageal cancer and has been used in clinical trials assessing the efficacy of neoadjuvant treatment [24].

The following four models were fitted, using the following sets of features:

1. 'Stage': overall TNM stage.
2. 'Clinical': age, clinical T-stage and N-stage, as determined at tumour board review
3. 'ClinVol': age, clinical T-stage and N-stage, primary CT tumour volume
4. 'ClinRad': age, clinical T-stage and N-stage, primary tumour volume, two GLCM features.

Discrimination and calibration of 3-year OS were assessed using the area under the receiver operator curve (AUC) and Brier score, respectively, using the riskRegression R library [25]. Confidence intervals were estimated using the method of Blanche [26] and compared using the Delong test. Following Van Rossum [27], clinical utility was assessed by thresholding the model prediction in training data to maximise specificity whilst maintaining sensitivity of >90%. Confidence intervals for sensitivity, specificity and accuracy were estimated with 1000 replacement bootstraps.

To ensure absence of dataset-partitioning bias or institutional confounding, supplementary post hoc model evaluation was performed using each institution in turn for testing and the remaining four institutions for model fitting. Kaplan–Meier curves were plotted, grouping participants according to the target 90% sensitivity threshold fitted in training data.

A post hoc analysis was also performed to estimate conditional dependencies between radiomic features and survival time, using data from training and testing cohorts. A partial Spearman correlation matrix was inferred using the de-sparsified graphical least absolute shrinkage and selection operator method via the SILGGM package and 95% confidence intervals were estimated via bootstrapping with 1000 replicates. Spearman correlation was also employed to assess volume confounding of radiomic features. Analysis of variance was employed to test radiomic stability with respect to scanner manufacturer and study institution.

Results

Participant and dataset characteristics

Of 210 participants recruited, 48 participants were excluded, generating a training set of 95 participants and a test set of 67 participants. The participant flowchart is provided in Fig. 1 showing the reasons for exclusions. Participant and dataset characteristics are shown in Table 1 and Supplementary Fig. 3.

Performance of model variables

Spearman correlations of the prespecified predictors in the training data are shown in Fig. 2A. GLCM_InverseDifferenceMoment was highly correlated with tumour volume ($r = +0.33$) and GLCM_Contrast ($r = -0.94$). Hence, GLCM_Correlation and GLCM_Contrast were selected for modelling (Supplementary Material). Clinical T-stage was the most important prognosticator, with each model assigning it a significant positive coefficient (representing increasing risk with increasing stage). The second most influential predictor was age, which was assigned significant positive coefficients in each model. Tumour volume was an insignificant predictor in both

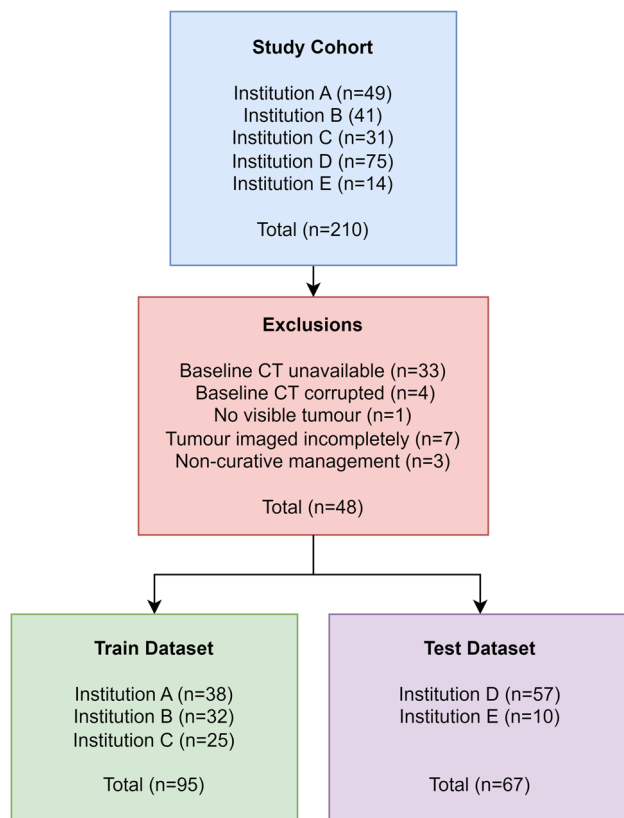


Fig. 1 Study participant flowchart

the ClinVol and ClinRad models. GLCM_Correlation was assigned a marginal negative coefficient in the ClinRad model, and GLCM_Contrast was the least influential model predictor. Model coefficients, their standard errors and associated z-tests are reported in Table 2. Histograms of model predictions demonstrated that each model had similar distributions of predictions in training and testing (Fig. 2B). Radiomic features are visualised at voxel level in Fig. 3.

Prediction of 3-year overall survival

The ClinRad model showed best discrimination of 3-year OS, achieving similar performance in both training (AUC, 0.71 [95% CI 0.60, 0.82]) and testing (AUC, 0.68 [95% CI 0.54, 0.81]) (Table 3). Test discrimination of the ClinRad model was greater than that of Stage alone (Δ AUC, 0.12 [95% CI 0.01, 0.23]; $p=0.04$). Stage was the least discriminative model in both training (AUC, 0.60 [95% CI 0.49, 0.71]) and testing (AUC, 0.56 [95% CI 0.44, 0.67]).

The Clinical model achieved similar test discrimination to the ClinRad model (AUC, 0.66 [95% CI 0.51, 0.80]; Δ AUC, 0.02 [95% CI -0.04, 0.08]; $p>0.05$). The ClinVol model attained marginally lower test discrimination than the Clinical model (AUC, 0.65 [95% CI 0.50, 0.79]). All models yielded similar test calibration. Only Stage and ClinRad models achieved target 90% sensitivity in both training and testing. Here, the ClinRad model yielded slightly higher specificity (specificity, 0.19 [95% CI 0.06,

Table 1 Participant characteristics in training and testing datasets

Variable	Value	Training	Test	p-value
Age (mean, SD)	Years	65 ± 9	65 ± 9	.89
Sex (n, %)	Male	11 (12)	7 (10)	
	Female	84 (88)	60 (90)	.99
Clinical T-stage (n, %)	T1	1 (1)	0 (0)	.45
	T2	16 (17)	11 (16)	
	T3	77 (81)	53 (79)	
	T4	1 (1)	3 (4)	
Clinical N-stage (n, %)	N0	24 (25)	13 (19)	.12
	N1	36 (38)	38 (57)	
	N2	32 (34)	14 (21)	
	N3	3 (3)	2 (3)	
Location (n, %)	Oesophagus, mid	5 (5)	0 (0)	< .001
	Oesophagus, lower	51 (54)	12 (18)	
	Gastrooesophageal junction	39 (41)	55 (82)	
Treatment (n, %)	Chemotherapy +/- radiotherapy and surgery	81 (85)	65 (97)	.005
	Chemoradiotherapy only	1 (1)	2 (3)	
	Surgery only	13 (14)	0 (0)	
Survival status at 3 years (n, %)	Deceased	57 (60)	30 (45)	.08
	Survived	38 (40)	37 (55)	

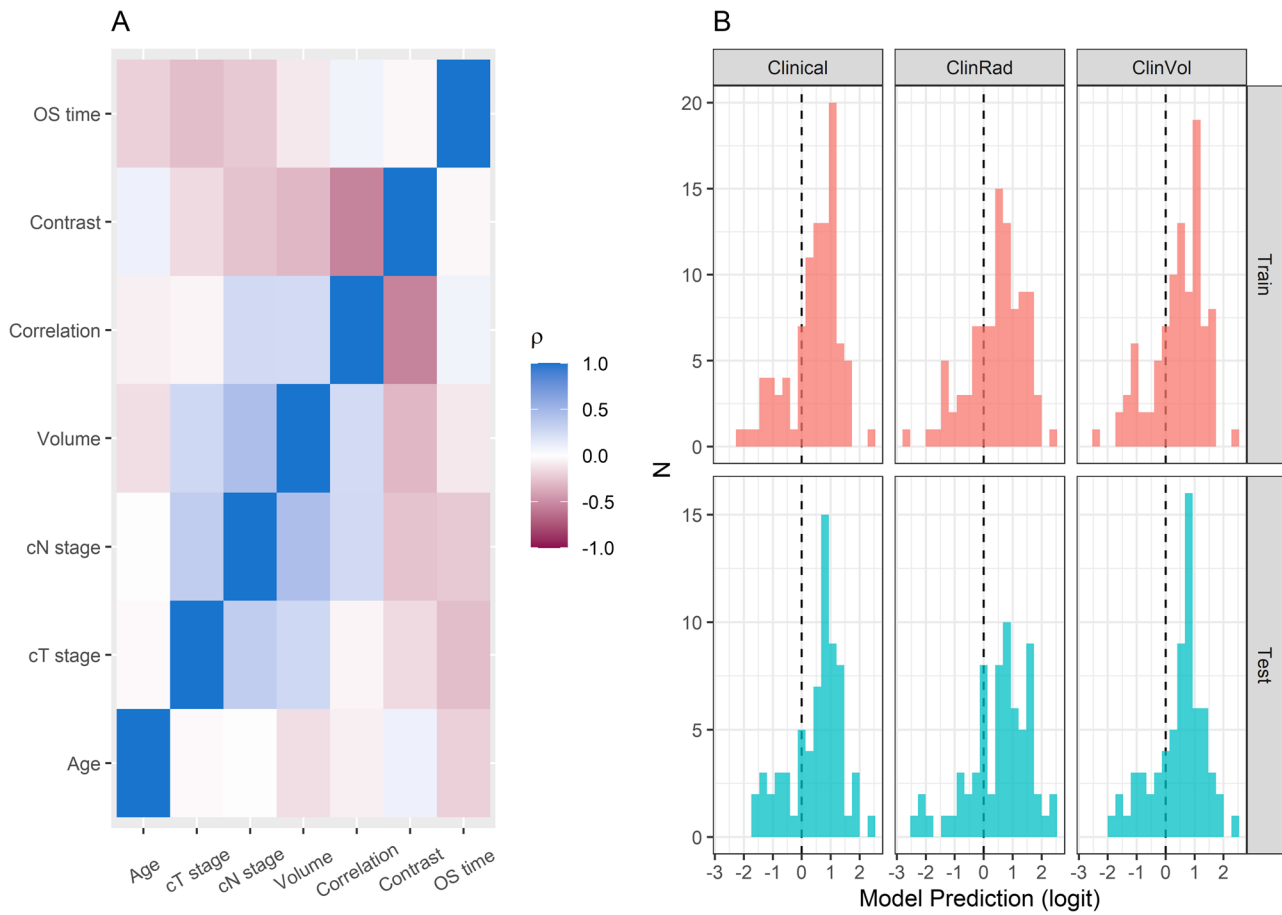


Fig. 2 Spearman correlations of predictor variables in training data (A) and histograms of model predictions in training and testing data (B) are shown

Table 2 Summary of model coefficients

Model	Variable	Coefficient	Standard error	p-value
Clinical	(Intercept)	-7.49	2.57	.004
	Age	0.05	0.03	.05
	cT-stage	1.45	0.61	.02
	cN-stage	0.32	0.30	.28
ClinVol	(Intercept)	-5.75	4.09	.16
	Age	0.05	0.03	.06
	cT-stage	1.50	0.62	.02
	cN-stage	0.39	0.32	.23
	Volume (log cm ³)	-0.18	0.33	.59
ClinRad	(Intercept)	-5.16	4.43	.24
	Age	0.05	0.03	.05
	cT-stage	1.41	0.64	.03
	cN-stage	0.48	0.33	.14
	Volume (log cm ³)	-0.10	0.34	.76
	GLCM_Correlation	-2.45	2.20	.26
	GLCM_Contrast	-0.03	0.32	.92

Logistic regression coefficients and p-values from unadjusted two-tailed z-tests are reported

0.33]) than Stage (specificity, 0.14 [95% CI 0.03, 0.26]) at this threshold.

Supplementary per-institution model testing results were consistent with the main external validation results (Supplementary Table 2). Kaplan–Meier curves are provided in Fig. 4. Survival statistics are provided in Supplementary Table 3. In the test data, risk groups assigned by the ClinRad model separated survival curves for the initial 3 years, and convergence was observed at 5 years. However, few participants were assigned to the high-risk group (9/67, 13%). Risk groupings assigned by Clinical and ClinVol models achieved little separation of survival curves in testing. Post hoc partial correlation analysis confirmed that, over both training and validation datasets, the most informative predictors of overall survival time were clinical N-stage (partial ρ , -0.15 [95% CI -0.33, 0.02]) and age (partial ρ , -0.14 [95% CI -0.3, 0.02]). The level of independent predictive information contributed by GLCM_Correlation (partial ρ , 0.10 [95% CI -0.05, 0.25]) was comparable to that of clinical T-stage (partial ρ , -0.12 [95% CI -0.26, 0.04]). Partial

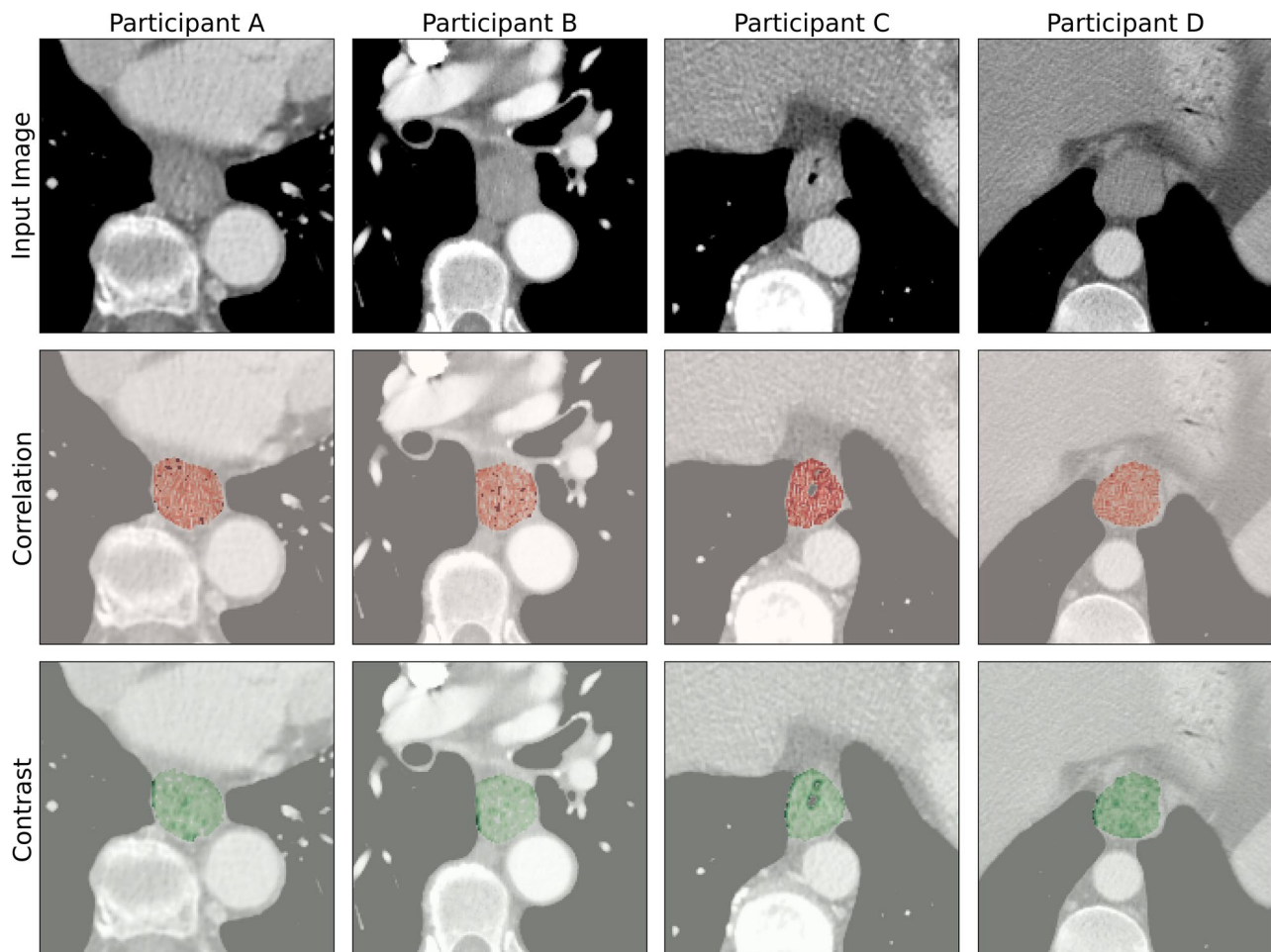


Fig. 3 Visualisation of CT images and voxel-level radiomic features (GLCM correlation and GLCM contrast) in four study participants. Participant **A**: 71-year-old female with a clinically staged T2 tumour [correlation 0.67, contrast 1.22]; participant **B**: 78-year-old female with a clinically staged T3 tumour [correlation 0.5, contrast 2.56]; participant **C**: 62-year-old male with a clinically staged T2 tumour [correlation 0.48, contrast 3.01]; participant **D**: 56-year-old male with a clinically staged T3 tumour [correlation 0.57, contrast 2.10]

correlation analysis results are provided in Supplementary Table 4. Feature variability with respect to scanner manufacturer and study institution are also presented in Supplementary Material.

Discussion

The ability to provide individualised risk–benefit analysis would help to optimise management decisions in potentially resectable oesophageal cancer. Improving prognostication is a step in this direction as prognosis influences treatment decisions made by doctors and patients. Surgery may have an impact on quality of life for up to 12 months post treatment [28] and patients with a poorer prognosis may not fully benefit from a multimodal approach. Initial imaging studies have suggested that radiomic features may have additive

prognostic value [11, 12, 18, 21]. However, prespecified models have to demonstrate reliable prediction in external datasets without local refitting. Accordingly, studies need to transition to the evaluation of previously proposed predictors and models, rather than continuing to fit new models with many degrees of freedom to new clinical data [18].

In this prospective multicentre study, we have demonstrated that a multivariate clinicoradiomic prognostic model (ClinRad) incorporating previously identified CT features improved discrimination of 3-year OS compared to TNM staging with an AUC of 0.68 in the test dataset, but offered similar calibration. The Clinical model had similar performance as the ClinRad model with an AUC of 0.66. Both Clinical and ClinRad models retained discriminative capacity

Table 3 Model discrimination and calibration metrics with respect to 3-year overall survival

Dataset	Model	AUC	Brier score	Sensitivity*	Specificity*	Accuracy*
Train	Stage	0.60 [0.49, 0.71]	0.24 [0.21, 0.28]	0.95 [0.89, 1.00]	0.21 [0.09, 0.35]	0.65 [0.56, 0.74]
Train	Clinical	0.69 [0.57, 0.80]	0.22 [0.18, 0.27]	0.91 [0.83, 0.98]	0.35 [0.21, 0.50]	0.69 [0.60, 0.78]
Train	ClinVol	0.70 [0.58, 0.81]	0.22 [0.18, 0.27]	0.91 [0.84, 0.98]	0.34 [0.20, 0.49]	0.68 [0.58, 0.77]
Train	ClinRad	0.71 [0.60, 0.82]	0.22 [0.17, 0.26]	0.93 [0.86, 0.98]	0.29 [0.14, 0.44]	0.67 [0.58, 0.77]
Test	Stage	0.56 [0.44, 0.67]	0.27 [0.24, 0.31]	0.93 [0.83, 1.00]	0.14 [0.03, 0.26]	0.49 [0.37, 0.61]
Test	Clinical	0.66 [0.51, 0.80]	0.28 [0.23, 0.33]	0.83 [0.68, 0.96]	0.22 [0.09, 0.36]	0.49 [0.37, 0.61]
Test	ClinVol	0.65 [0.50, 0.79]	0.28 [0.23, 0.33]	0.83 [0.69, 0.96]	0.19 [0.07, 0.32]	0.48 [0.36, 0.60]
Test	ClinRad	0.68 [0.54, 0.81]	0.28 [0.22, 0.33]	0.90 [0.77, 1.00]	0.19 [0.06, 0.33]	0.51 [0.39, 0.63]

*Sensitivity, specificity and accuracy were measured at a risk threshold which achieved at least 90% sensitivity in training data. 95% confidence intervals are provided in square brackets

between training and testing, though calibration deteriorated, suggesting a distributional mean shift between institutions.

Our findings are concordant with previously published data of Larue et al [12], where the high-dimensional random forest radiomic model with other features achieved AUCs of 0.69 and 0.61 in training and testing, respectively. The direction of radiomic coefficients fitted in this study is consistent with previously published results by Zhang et al [21], who observed increasing GLCM_Correlation in patients with oesophageal adenocarcinoma who responded to chemoradiotherapy. In our model, low GLCM_Correlation was an adverse prognosticator. Our finding that GLCM correlation was the most informative predictor also concurs with Klaasen et al [22]. However, as Klaasen's model employed a random forest architecture, the directional concordance of results could not be verified.

Zhang et al [21] also observed decreasing GLCM contrast in chemoradiotherapy responders. In our study, GLCM contrast did not affect model predictions substantially, indicating that any prognostic information it encoded was already provided by the other clinical and image-based predictors already modelled.

An advantage of our study is that it incorporated multicentre prospective data, thereby providing realistic conditions for the estimation of model informativeness and generalisability. The imaging equipment and protocols were representative of the varying conditions

encountered in clinical practice. The imaging acquisition parameters in this dataset reflected typical clinical practice and variations between institutions, which a radiomic model must be able to accommodate. We noted that GLCM correlation and GLCM contrast varied according to institution and scanner manufacturer respectively. This variability introduces noise which can complicate modelling of the underlying prognostic signal. Clinical deployment of radiomic models requires either that this noise is accommodated or that clinical imaging protocols adapt to acquire images under more standard conditions.

Model validation was performed in test data from three institutions which were unobserved during model development, yielding a realistic estimate of model generalisability in our healthcare system. However, our study had limitations. First, manual segmentation especially of early-stage cancers is subject to intra-reader and inter-reader variability [29]. Second, radiomic approaches are not typically well suited for the identification of new imaging biomarkers, due to the low ratio of events to evaluated variables [18, 20]. It is noteworthy that the ClinRad model fitted here is simpler than that of Larue, whilst matching its training performance, and marginally improving upon its generalisation [12]. However, a necessary cost of this study design is that the other informative radiomic features may have been omitted. Third, although both the ClinRad model and TNM staging demonstrated 90% sensitivity, the low specificity achieved at this threshold is a limitation. Fourth, the

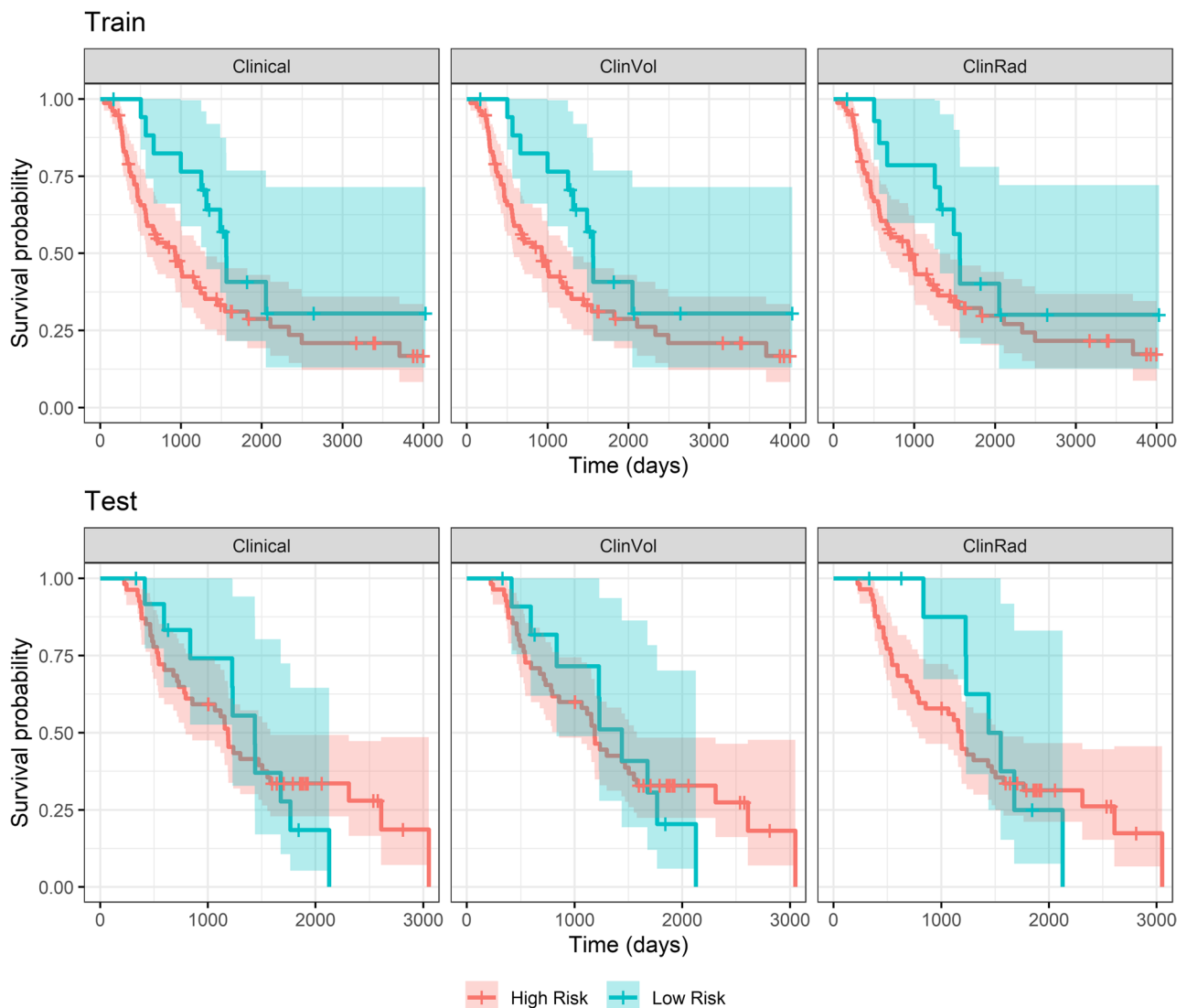


Fig. 4 Kaplan–Meier plot of survival in high- and low-risk groups according to each model score. Risk groups were defined according to the target 90% sensitivity threshold fitted in training data

improvement in performance between the Clinical and ClinRad model is small and unlikely to change clinical management substantially. Finally, the logistic regression models employed in this analysis were insensitive to non-linear and nonmonotonic effects.

In conclusion, we have confirmed in a prospective multicentre dataset that previously proposed GLCM features—correlation and contrast—contain incremental prognostic information. The clinicoradiomic model incorporating GLCM correlation and contrast with tumour and nodal stage, age and volume outperformed TNM stage alone in the discrimination of 3-year overall survival. Nevertheless, the level of discrimination

remained modest and it is questioned if this will impact on management substantially.

Abbreviations

AUC	Area under receiver operating characteristic curve
GLCM	Grey-level co-occurrence matrix
OS	Overall survival
TNM	Tumour-node-metastasis

Supplementary information

The online version contains supplementary material available at <https://doi.org/10.1007/s00330-024-10666-y>.

Below is the link to the electronic supplementary material. Supplementary file1 (PDF 544 KB)

Acknowledgements

This manuscript was submitted on behalf of the OCCAMS Consortium. Oesophageal Cancer Clinical and Molecular Stratification (OCCAMS) Consortium:

Rebecca C. Fitzgerald¹, Paul A.W. Edwards^{1,2}, Nicola Grehan^{1,5}, Barbara Nutzinger¹, Aisling M Redmond¹, Sujath Abbas¹, Adam Freeman¹, Elizabeth C. Smyth³, Maria O'Donovan^{1,3}, Ahmad Miremadi^{1,3}, Shalini Malhotra^{1,3}, Monika Tripathi^{1,3}, Calvin Cheah¹, Hannah Coles¹, Matthew Eldridge², Maria Secrier², Ginny Devonshire², Sriganesh Jammula², Jim Davies⁴, Charles Crichton⁴, Nick Carroll⁵, Richard H.Hardwick⁵, Peter Safranek⁵, Andrew Hindmarsh⁵, Vijayendran Sujendran⁵, Stephen J. Hayes^{6,13}, Yeng Ang^{6,7,26}, Andrew Sharrocks²⁶, Shaun R. Preston⁸, Izhar Bagwan⁸, Vicki Save⁹, J. Robert O'Neill^{5,9,20}, Olga Tucker^{10,29}, Andrew Beggs^{10,25}, Philippe Taniere¹⁰, Sonia Puig¹⁰, Gianmarco Contino¹⁰, Timothy J. Underwood^{11,12}, Ben L. Grace¹¹, Jesper Lagergren^{14,22}, Andrew Davies^{14,21}, Fujun Chang^{14,21}, Ula Mahadeva¹⁴, Francesca D. Ciccarelli²¹, Grant Sanders¹⁵, David Chan¹⁵, Ed Cheong¹⁶, Bhaskar Kumar¹⁶, Loveena Sreedharan¹⁶, Irshad Soomro¹⁷, Philip Kaye¹⁷, John Saunders^{6,17}, Laurence Lovat¹⁸, Rehan Haidry¹⁸, Michael Scott¹⁹, Sharmila Sothi²³, George B. Hanna²⁷, Christopher J. Peters²⁷, Krishna Moorthy²⁷, Anna Grabowska²⁸, Richard Turkington³⁰, Damian McManus³⁰, Helen Coleman³⁰, Russell D Petty³¹, Freddie Bartlett³², Tom D.L. Crosby³³

¹Early Cancer Institute, Department of Oncology, Hutchison Research Centre, University of Cambridge, Cambridge, CB2 0XZ, UK
²Cancer Research UK Cambridge Institute, University of Cambridge, Cambridge, CB2 0RE, UK
³Department of Histopathology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, CB2 0QQ, UK
⁴Department of Computer Science, University of Oxford, Oxford, OX1 3QD, UK
⁵Cambridge University Hospitals NHS Foundation Trust, Cambridge, CB2 0QQ, UK
⁶Salford Royal NHS Foundation Trust, Salford, M6 8HD, UK
⁷Wigan and Leigh NHS Foundation Trust, Wigan, Manchester, WN1 2NN, UK
⁸Royal Surrey County Hospital NHS Foundation Trust, Guildford, GU2 7XX, UK
⁹Edinburgh Royal Infirmary, Edinburgh, EH16 4SA, UK
¹⁰University Hospitals Birmingham NHS Foundation Trust, Birmingham, B15 2GW, UK
¹¹University Hospital Southampton NHS Foundation Trust, Southampton, SO16 6YD, UK
¹²Cancer Sciences Division, University of Southampton, Southampton, SO17 1BJ, UK
¹³Faculty of Medical and Human Sciences, University of Manchester, Manchester, M13 9PL, UK
¹⁴Guy's and St Thomas's NHS Foundation Trust, London, SE1 7EH, UK
¹⁵Plymouth Hospitals NHS Trust, Plymouth, PL6 8DH, UK
¹⁶Norfolk and Norwich University Hospital NHS Foundation Trust, Norwich, NR4 7UY, UK
¹⁷Nottingham University Hospitals NHS Trust, Nottingham, NG7 2UH, UK
¹⁸University College London, London, WC1E 6BT, UK
¹⁹Wythenshawe Hospital, Manchester, M23 9LT, UK
²⁰Edinburgh University, Edinburgh, EH8 9YL, UK
²¹King's College London, London, WC2R 2LS, UK
²²Karolinska Institute, Stockholm, SE-171 77, Sweden
²³University Hospitals Coventry and Warwickshire NHS, Trust, Coventry, CV2 2DX, UK
²⁴Peterborough Hospitals NHS Trust, Peterborough City Hospital, Peterborough, PE3 9GZ, UK
²⁵Institute of Cancer and Genomic Sciences, University of Birmingham, Birmingham, B15 2TT, UK
²⁶Gastrointestinal Science Centre, University of Manchester, M13 9PL, UK
²⁷Department of Surgery and Cancer, Imperial College, London, W2 1NY, UK
²⁸Queen's Medical Centre, University of Nottingham, Nottingham, NG7 2UH, UK
²⁹Heart of England NHS Foundation Trust, Birmingham, B9 5SS, UK
³⁰Centre for Cancer Research and Cell Biology, Queen's University Belfast, Belfast, BT7 1NN, Northern Ireland
³¹Tayside Cancer Centre, Ninewells Hospital and Medical School, Dundee, DD1 9SY, Scotland, UK
³²Portsmouth Hospitals NHS Trust, Portsmouth, PO6 3LY, UK
³³Velindre University NHS Trust, Cardiff, Wales, CF15 7QZ, UK

Funding

The authors acknowledge funding support from the following: UK Research & Innovation London Medical Imaging and Artificial Intelligence Centre; Wellcome/Engineering and Physical Sciences Research Council Centre for Medical Engineering at King's College London [WT 203148/Z/16/Z]; National Institute for Health Research Biomedical Research Centre at Guy's & St Thomas' Hospitals and King's College London; National Institute for Health Research Biomedical Research Centre at Guy's & St Thomas' Hospitals and King's College London; Cancer Research UK National Cancer Imaging Translational Accelerator [C1519/A28682]. OCCAMS was funded by a Programme Grant from Cancer Research UK (RG66287, A15874). OCCAMS2 was funded by a Programme Grant from Cancer Research UK (RG81771/RG84119, A22720/A22131).

Declarations

Guarantor

The scientific guarantor of this publication is Professor Vicky Goh.

Conflict of interest

The authors of this manuscript declare no relationships with any companies, whose products or services may be related to the subject matter of the article.

Statistics and biometry

One of the authors has significant statistical expertise. No complex statistical methods were necessary for this paper.

Informed consent

Written informed consent was obtained from all subjects (patients) in this study.

Ethical approval

Institutional Review Board approval was obtained.

Study subjects or cohorts overlap

None of these study subjects or cohorts has been previously reported.

Methodology

- prospective
- diagnostic or prognostic study
- multicentre study

Author details

¹Department of Cancer Imaging, School of Biomedical Engineering and Imaging Sciences, King's College London, London, UK. ²Department of Radiology, Royal Marsden Hospital NHS Trust, Sutton, Surrey, UK. ³Department of Clinical Oncology, Guy's & St Thomas' Hospitals NHS Foundation Trust, London, UK. ⁴Department of Surgery, Guy's & St Thomas' Hospitals NHS Foundation Trust, London, UK. ⁵Department of Radiology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. ⁶Department of Radiology, Nottingham University Hospitals NHS Foundation Trust, Nottingham, UK. ⁷Department of Surgery, Nottingham University Hospitals NHS Foundation Trust, Nottingham, UK. ⁸Department of Surgery, Edinburgh Royal Infirmary, NHS Lothian, Edinburgh, UK. ⁹Department of Radiology, Guy's & St Thomas' Hospitals NHS Foundation Trust, Westminster Bridge Road, London SE1 7EG, UK.

Received: 29 August 2023 Revised: 29 January 2024

Accepted: 6 February 2024 Published online: 25 March 2024

References

1. Arnold M, Ferlay J, van Berge Henegouwen MI, Soerjomataram I (2020) Global burden of oesophageal and gastric cancer by histology and subsite in 2018. *Gut* 69:1564–1571
2. Cunningham D, Allum WH, Stenning SP et al (2006) Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med* 355:11–20

3. Medical Research Council Oesophageal Cancer Working Group (2002) Surgical resection with or without preoperative chemotherapy in oesophageal cancer: a randomised controlled trial. *Lancet* 359:1727–1733
4. van Hagen P, Hulshof MC, van Lanschot JJ et al (2012) Preoperative chemoradiotherapy for esophageal or junctional cancer. *N Engl J Med* 366:2074–2084
5. Morgan E, Soerjomataram I, Gavin AT et al (2021) International trends in oesophageal cancer survival by histological subtype between 1995 and 2014. *Gut* 70:234–242
6. Sudo K, Xiao L, Wadhwa R et al (2014) Importance of surveillance and success of salvage strategies after definitive chemoradiation in patients with esophageal cancer. *J Clin Oncol* 32:3400–3405
7. Obermannová R, Alsina M, Cervantes A et al (2022) Oesophageal cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann Oncol* 33:992–1004
8. Wolfson P, Ho KMA, Bassett P et al (2021) Accuracy of clinical staging for T2N0 oesophageal cancer: systematic review and meta-analysis. *Dis Esophagus* 34:1–12
9. Tariman JD, Doorenbos A, Schepp KG, Singhal S, Berry DL (2014) Information needs priorities in patients diagnosed with cancer: a systematic review. *J Adv Pract Oncol* 2014:115–122
10. Foley KG, Hills RK, Berthon B et al (2018) Development and validation of a prognostic model incorporating texture analysis derived from standardised segmentation of PET in patients with oesophageal cancer. *Eur Radiol* 28:428–436
11. Piazzese C, Foley K, Whybra P, Hurt C, Crosby T, Spezi E (2019) Discovery of stable and prognostic CT-based radiomic features independent of contrast administration and dimensionality in oesophageal cancer. *PLoS One* 14:e0225550
12. Larue R, Klaassen R, Jochems A et al (2018) Pre-treatment CT radiomics to predict 3-year overall survival following chemoradiotherapy of esophageal cancer. *Acta Oncol* 57:1475–1481
13. Zwanenburg A, Vallières M, Abdalah MA et al (2020) The Image Biomarker Standardization Initiative: standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* 295:328–338
14. Mason DL (2011) Pydicom: an open source DICOM library. *Med Phys* 38, 6(10):3493
15. Umeoka S, Koyama T, Togashi K et al (2006) Esophageal cancer: evaluation with triple-phase dynamic CT—initial experience. *Radiology* 239:777–783
16. Umeoka S, Koyama T, Watanabe G et al (2010) Preoperative local staging of esophageal carcinoma using dual-phase contrast-enhanced imaging with multi-detector row computed tomography: value of the arterial phase images. *J Comput Assist Tomogr* 34:406–412
17. van Griethuysen JJM, Fedorov A, Parmar C et al (2017) Computational radiomics system to decode the radiographic phenotype. *Cancer Res* 77:e104–e107
18. O'Shea RJ, Rookyard C, Withey S, Cook GJR, Tsoka S, Goh V (2022) Radiomic assessment of oesophageal adenocarcinoma: a critical review of 18F-FDG PET/CT, PET/MRI and CT Insights Imaging 13:104
19. Moons KG, Altman DG, Reitsma JB et al (2015) Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 162:W1–73
20. Lambin P, Leijenaar RTH, Deist TM et al (2017) Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol* 14:749–762
21. Zhang YH, Herlin G, Rouvelas I, Nilsson M, Lundell L, Brismar TB (2018) Texture analysis of computed tomography data using morphologic and metabolic delineation of esophageal cancer—relation to tumor type and neoadjuvant therapy response. *Dis Esophagus* 32:1–10
22. Klaassen R, Larue R, Mearadji B et al (2018) Feasibility of CT radiomics to predict treatment response of individual liver metastases in esophago-gastric cancer patients. *PLoS One* 13:e0207362
23. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 49:1373–1379
24. Shapiro J, van Lanschot JJB, Hulshof M et al (2015) Neoadjuvant chemoradiotherapy plus surgery versus surgery alone for oesophageal or junctional cancer (CROSS): long-term results of a randomised controlled trial. *Lancet Oncol* 16:1090–1098
25. Gerds AT, Ozenne B (2020) riskRegression: risk regression models and prediction scores for survival analysis with competing risks. R package version 5
26. Blanche P, Dartigues JF, Jacqmin-Gadda H (2013) Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 32:5381–5397
27. van Rossum PS, Fried DV, Zhang L et al (2016) The incremental value of subjective and quantitative assessment of 18F-FDG PET for the prediction of pathologic complete response to preoperative chemoradiotherapy in esophageal cancer. *J Nucl Med* 57:691–700
28. Lagergren P, Avery KN, Hughes R et al (2007) Health-related quality of life among patients cured by surgery for esophageal cancer. *Cancer* 110:686–693
29. Li F, Li Y, Wang X et al (2022) Inter-observer and intra-observer variability in gross tumor volume delineation of primary esophageal carcinomas based on different combinations of diagnostic multimodal images. *Front Oncol* 12:817413

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.