Prognostic Modeling of Patients Undergoing Surgery Alone for Esophageal Squamous Cell Carcinoma: A Histopathological and Computed Tomography Based Quantitative Analysis

Lei-Lei Wu [1‡], Jin-Long Wang [2‡], Wei Huang [1‡], Xuan Liu [1‡], Yang-Yu Huang [1], Jing Zeng [1], Chun-Yan Cui [1], Jia-Bin Lu [1], Peng Lin [1], Hao Long [1], Lan-Jun Zhang [1], Jun Wei [3†§], Yao Lu [2§] and Guo-Wei Ma [1*§]

[1] Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University, Guangzhou, China, [2] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China, [3] Department of Radiology, University of Michigan, Ann Arbor, MI, United States

Objective: To evaluate the effectiveness of a novel computerized quantitative analysis based on histopathological and computed tomography (CT) images for predicting the postoperative prognosis of esophageal squamous cell carcinoma (ESCC) patients.

Methods: We retrospectively reviewed the medical records of 153 ESCC patients who underwent esophagectomy alone and quantitatively analyzed digital histological specimens and diagnostic CT images. We cut pathological images (6000 × 6000) into 50 × 50 patches; each patient had 14,400 patches. Cluster analysis was used to process these patches. We used the pathological clusters to all patches ratio (PCPR) of each case for pathological features and we obtained 20 PCPR quantitative features. Totally, 125 computerized quantitative (20 PCPR and 105 CT) features were extracted. We used a recursive feature elimination approach to select features. A Cox hazard model with L1 penalization was used for prognostic indexing. We compared the following prognostic models: Model A: clinical features; Model B: quantitative CT and clinical features; Model C: quantitative histopathological and clinical features; and Model D: combined information of clinical, CT, and histopathology. Indices of concordance (C-index) and leave-one-out cross-validation (LOOCV) were used to assess prognostic model accuracy.

Results: Five PCPR and eight CT features were treated as significant indicators in ESCC prognosis. C-indices adjusted for LOOCV were comparable among four models, 0.596 (Model A) vs. 0.658 (Model B) vs. 0.651 (Model C), and improved to 0.711with Model D combining information of clinical, CT, and histopathology (all p<0.05). Using Model D, we stratified patients into low- and high-risk groups. The 3-year overall survival rates of low- and high-risk patients were 38.0% and 25.0%, respectively (p<0.001).

Conclusion: Quantitative prognostic modeling using a combination of clinical data, histopathological, and CT images can stratify ESCC patients with surgery alone into high-risk and low-risk groups.

Keywords: quantitative analysis, esophageal cancer, prognosis, medical images, survival model

INTRODUCTION

Esophageal cancer (EC) is the seventh most common cancer and the sixth leading cause of cancer-related mortality globally (1). The main histological subtypes of EC include esophageal squamous cell carcinoma (ESCC) and esophageal adenocarcinoma (EA) (1–3). ESCC accounts for approximately 79% of ECs worldwide, and in China, more than 90% of all ECs are ESCCs (4, 5). Esophagectomy is a routine treatment for ESCC in current clinical practice. However, the 5-year overall survival rate of ESCC is only 20–40% (6) and the postoperative care of patients with ESCC remains challenging. Therefore, the current postoperative care protocol for the management of patients with ESCC needs to be optimized to improve their life expectancy.

An important aspect of postoperative care is the development of tools for disease staging and prognosis prediction, where the goal is to assist physicians in the pre-selection of patients with poor prognosis who may have a high risk of metastasis and/or recurrence. If successful, such tools will improve the curative rates for ESCC. The available studies suggest that radiological image analysis plays an important role in the prognosis of patients with malignant tumors. Previous studies demonstrated the role of quantitative analysis of computed tomography (CT) and pathological images in prognosis prediction for lung cancer (7–9), hypopharyngeal cancer (10), colorectal cancer (11), and other such cancers (12–14). In very few studies based on immunohistochemical staining for the diagnosis of the condition and analysis of disease prognosis (15–19), pathological hematoxylin-eosin (HE) staining images were used for direct feature extraction and prognosis analysis (20–22). This is because the delineation of pathological images is time-consuming, and delineating tumor cells and computerizing the morphology of tumor cells. Recently, for pathological quantitative analysis, some studies used deep learning systems to elevate the accuracy of diagnosis in prostate cancer and improve the predictive ability for the prognosis of malignant tumors (21, 22).

In this study, we aimed to develop a prognostic prediction model of ESCC using CT, pathological, and clinical data. We extracted the features from the diagnostic CT images and pathological images of patients with ESCC, from which the useful information was extracted using a clustering analysis method. This is the novel report describing the development of a prognostic prediction model using the clustering analysis method for extracting useful artifacts from collected pathological data. The schematic diagram of the model is illustrated in **Figure 1A**.

METHODS

Study Cohort

The Ethics Committee of Sun Yat-sen University Cancer Center approved the study's protocol (Approval No: GZR2019-161), the requirement of informed consent was waived. The eligible cohort included ESCC patients with pathological and CT findings who underwent esophagectomy at the Department of Thoracic Surgery of a university cancer center in China between January 2014 and August 2015. Patients were excluded if they: (1) had received neoadjuvant chemotherapy, radiotherapy, or immunotherapy; (2) had been diagnosed with another malignancy currently or in the past; (3) had incomplete resection or residual tumor cells; (4) had died due to postoperative complications; (5) had tumors located at the esophagogastric junction or cervical esophagus; or (6) had other histological subtypes of esophageal cancer besides ESCC. According to the patient records, we translated the pathological staging according to the 8th edition of the American Joint Committee on Cancer (AJCC) Staging Manual. Under our study's protocol, a total of 153 patients were retrospectively analyzed.

Surgery

Popular esophagectomy procedures for ESCC include the Sweet (left thoracotomy and diaphragm incision), the McKeown (right thoracotomy, laparotomy, and neck incision), and the Ivor Lewis (laparotomy and right thoracotomy). In this cohort of patients, thoracoabdominal lymphadenectomy was performed as the standard procedure. During the operation, the median and mean the number of lymph nodes dissected was 18.0 and 20.54, respectively.

Follow-Up

We recommended patients visit the outpatient department for follow-up examinations every 3 months for the first 2 years, then every 6 months for the next 3 years, and annually thereafter. Follow-up examinations consisted of recording history, physical examination, chest radiography, cervical ultrasonography, abdominal ultrasonography, barium esophagography, and neck-abdomen CT scans. If necessary, patients also underwent positron emission tomography-CT and/or endoscopy examinations. In the study cohort, the median follow-up time was 28.0 months.

Tissue Microarray (TMA)

The use of the ESCC TMA was authorized by the ethics committee of Sun Yat-sen University Cancer Center (Approval No: GZR2019-161). The pathologist first reviewed the pathological HE staining slides and sketched the most
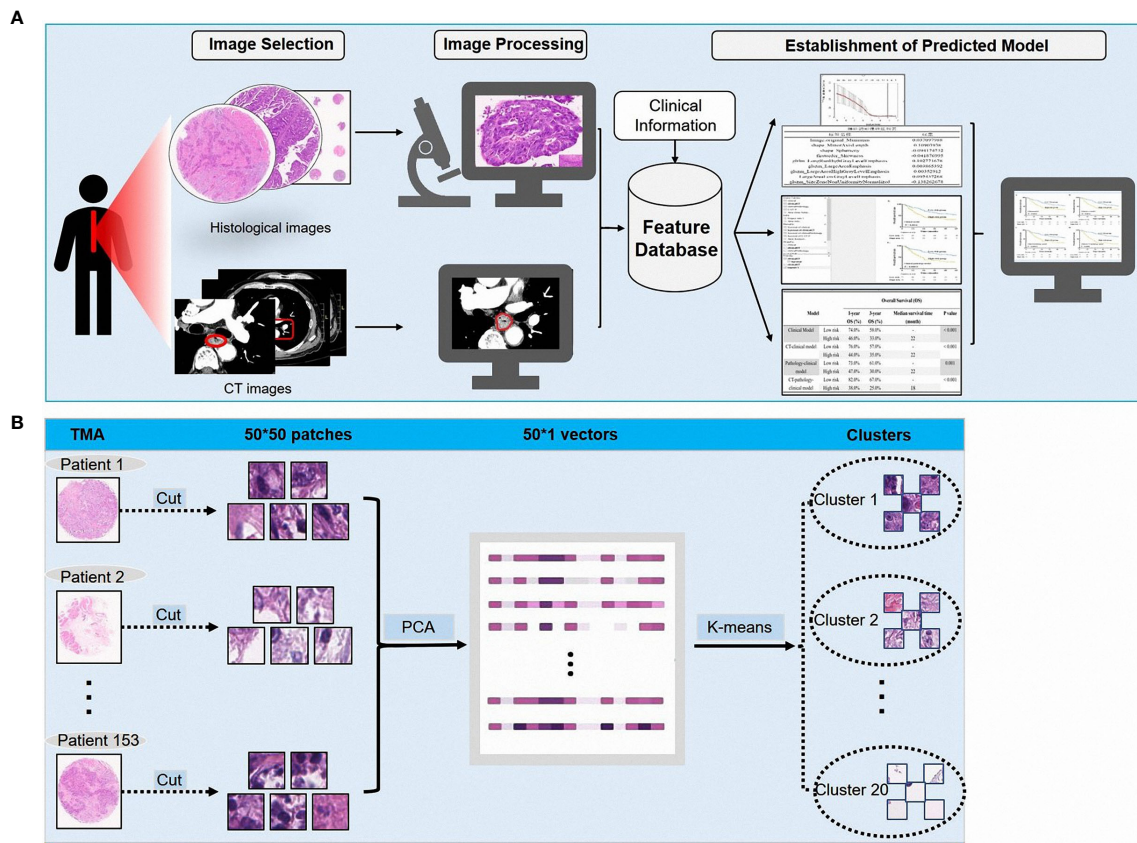
FIGURE 1 | **(A)** The schematic diagram showing the study design; **(B)** The processing of pathological images.

abundant areas of the malignant tumor. Then, the corresponding positions of the sketched areas were punched into the paraffin tissue. HE staining was performed to form a tissue microarray. We performed whole-slide microarray imaging to obtain an integral pathological image.

Quantitative Imaging Analysis

The radiologist sketched the 3D regions of interest (ROIs) of all images obtained on CT using an open-source software, Amira 5.4 (Thermo Fisher Scientific, Inc., Waltham, Massachusetts, USA). We extracted CT features from 3D CT ROIs; 105 CT quantitative features were extracted by using Pyradiomics 2.2.0 (in Python software 3.6; https://www.python.org).

Pathological images represent biological information at a different scale from that of CT images. Previous studies focused more on gray level features, geometric features, and texture features extracted from CT images. For pathological images, image patterns represent the internal characteristics and the external distribution of cells particularly related to clinical outcomes. Given the cellular level of the pathological scale, to utilize the histological information and further explore the pathological patterns, we cut the images (6000 × 6000) of each patient into 50 × 50 patches, and obtained a total of 14,400 patches for every patient. This size of patches was convenient to

observe the pattern of ESCCs and to further analyze the medical images. To determine the distribution of these patches in patients with ESCC, we used a cluster analysis method. A principal component analysis (PCA) and a K-means algorithm were used to divide all patches into 20 clusters, with each cluster representing one kind of cell-scale patch. In the pathological slide of each patient, the patch ratio of each cluster represented the ratio of the kind of cells. We used the pathological patches cluster ratio (PPCR) for pathological features and we obtained 20 PPCR quantitative features. **Figure 1A** illustrates the image analysis pipeline. **Figure 1B** shows an example of a cell level patch from the whole microarray slide.

Statistical Analysis

Statistical analyses were performed using SPSS Statistics 25.0 software (IBM SPSS, Inc., Chicago, IL, USA) and R version 3.6.2 (https://www.r-project.org/). Hazard ratios (HR) with 95% confidence intervals (95% CIs) were calculated by Cox univariate and multivariate regression analyses methods. Heatmaps were used to present with the distribution of 20 PPCR features in 153 patients with ESCC after cluster analysis. Recursive feature elimination (RFE) was used to select features from clinical features, CT features, and pathological features (23). The factors with a two-sided p-value of <0.2 were

incorporated into the constructing model. We used an L1-norm penalized Cox (LASSO-Cox) model, which could effectively prevent overfitting in the multivariate survival analysis, and found the best-penalized coefficient each time. The clinical survival model contained four features, the CT survival model contained eight CT quantitative features, the pathological survival model contained five quantitative PPCR features, and the CT-pathology-clinical model used all 17 quantitative features.

We built four LASSO-Cox models, combined all features into a survival score (SC), and divided patients into two groups (high-risk group and low-risk group) according to the median value of the SC. Harrell's C index was generated for discrimination of the prognostic model. To verify the stability of the small sample model, we performed leave-one-out cross-validations (LOOCV). According to the LOOCV results, we found that the improvement was significant when we combined the medical image features with clinical features to build a LASSO-Cox model. Kaplan–Meier analysis and log-rank tests were used to compare survival curves between groups. A two-sided p-value of <0.05 was considered statistically significant. Cases were censored at death or the end of follow-up. The selection of overall survival as a primary clinical endpoint was considered most clinically relevant.

RESULTS

Clinical Features

The clinical characteristics of patients in the cohort are listed in **Table 1**. Of the 153 enrolled patients, 123 (80.4%) were males and 30 (19.6%) were females, with an age range from 41 years to 84 years (median, 61 years). Tumor differentiation was well-moderated in most cases (N = 130, 85.0%) and poor in the rest of the patients (N = 23, 15.0%). The most common tumor, node, metastasis (TNM) stages of patients were IB (N = 56, 36.6%) and IIIA (N = 64, 41.8%). Ninety-one patients (59.5%) had a smoking history and 64 patients (41.8%) had a history of alcohol consumption. The 1- and 3-year overall survival rates were 60.0% and 46.0%, respectively. The median time from the operation to the last censoring date was 28.0 months, and the mean time was 31.84 months.

Extraction and Selection of Features of Images

A total of 105 textural features (six types of radiomics features, **Table 2**) were extracted from CT images from the 153 patients in the cohort. Following the extraction of pathological features from the pathological TMA images, we got 20 PPCR features. In the heatmap, we could observe the distribution of all clusters in 153 patients with ESCC, and Cluster 10 presented with a color gradient process (**Figure 2**). We found that Cluster 3 had the highest ratio in all patches (**Figure 3A**). Then we used RFE to select a total of four clinical features, and 13 medical image

TABLE 1 | The characteristic of clinical information in patients with ESCC.

Variable	No.	%
Sex		
Male	123	80.4
Female	30	19.6
Age (years)		
≤65	111	72.5
>65	42	27.5
Smoking status		
No	62	40.5
Yes	91	59.5
Drinking status		
No	89	58.2
Yes	64	41.8
Tumor location		
Upper	3	2.0
Middle	70	45.8
Lower	80	52.3
Tumor length (cm)		
≤5	115	75.2
>5	38	24.8
Differentiation		
Well–moderate	130	85.0
Poor	23	15.0
TNM stage		
IB	5	3.3
IIA	56	36.6
IIB	8	5.2
IIIA	8	5.2
IIIB	64	41.8
IV	12	7.8
T stage		
T1	10	6.5
T2	32	21.0
T3	107	69.9
T4a	4	2.6
N stage		
N0	69	45.1
N1	45	29.4
N2	28	18.3
N3	11	7.2
Nerve invasion		
No	81	52.9
Yes	72	47.1
Vascular invasion		
No	147	96.1
Yes	6	3.9
Amount of bleeding		
≤ 400ml	152	99.3
> 400ml	1	0.7
CRP		
≤5mg/ml	106	69.3
>5mg/ml	47	30.7
AGR		
≤ 1.5	96	62.7
> 1.5	57	37.3
Tumor thrombus		
No	100	65.4
Yes	53	34.5
Adjuvant therapy		
No	146	95.4
Chemotherapy	3	2.0
Radiotherapy	4	2.6

features which included five PPCR features and eight CT features (**Table 3** and **Figure 3B**).

TABLE 2 | Six types of radiomics features and the number of them.

Feature class	Count
Shape 3D	16
First Order Statistics	19
Gray Level Co-occurrence Matrix	24
Gray Level Run Length Matrix	16
Gray Level Size Zone Matrix	16
Gray Level Dependence Matrix	14

Construction of Survival Prediction Model

After constructing the survival models to predict the prognosis of ESCC and the determination of the best-penalized coefficients with the LASSO-Cox model, four survival models were created. The clinical survival model included the TNM stage, the concentration of C-reactive protein (CRP), the albumin-globulin ratio (AGR), and the history of alcohol consumption. The CT-clinical survival model contained eight CT features and four clinical features. The pathology-clinical survival model contained five PPCR features and four clinical variates. The CT-pathology-clinical model used all 17 features. The LOOCV-based C-indexes for the clinical survival model, CT-clinical survival model, pathology-clinical model, and CT-pathology-clinical survival model were 0.596, 0.658, 0.651, and 0.711, respectively. The predictive ability of the CT-pathology-clinical survival model was better than others (all P < 0.05). We found that the improvement was significant when we combined the pathological features with CT features to build a LASSO-Cox model.
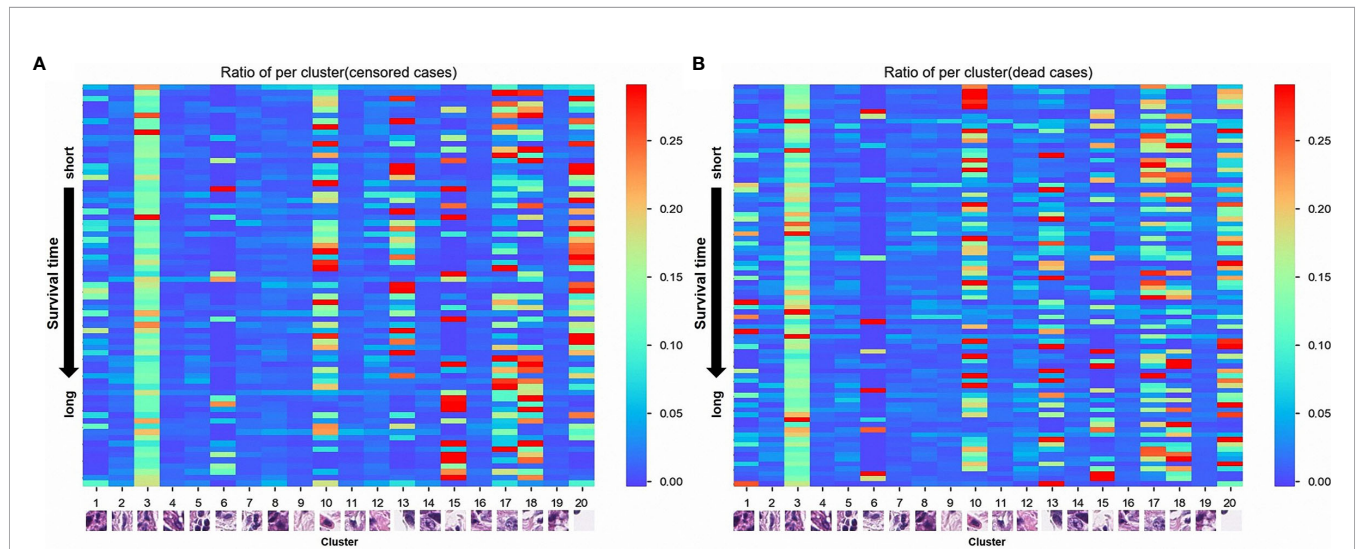

FIGURE 2 | (A) The distribution of 20 pathological patches cluster ratio (PPCR) features in alive patients with esophageal squamous cell cancer (ESCC); **(B)** The distribution of 20 PPCR features in dead patients with ESCC.
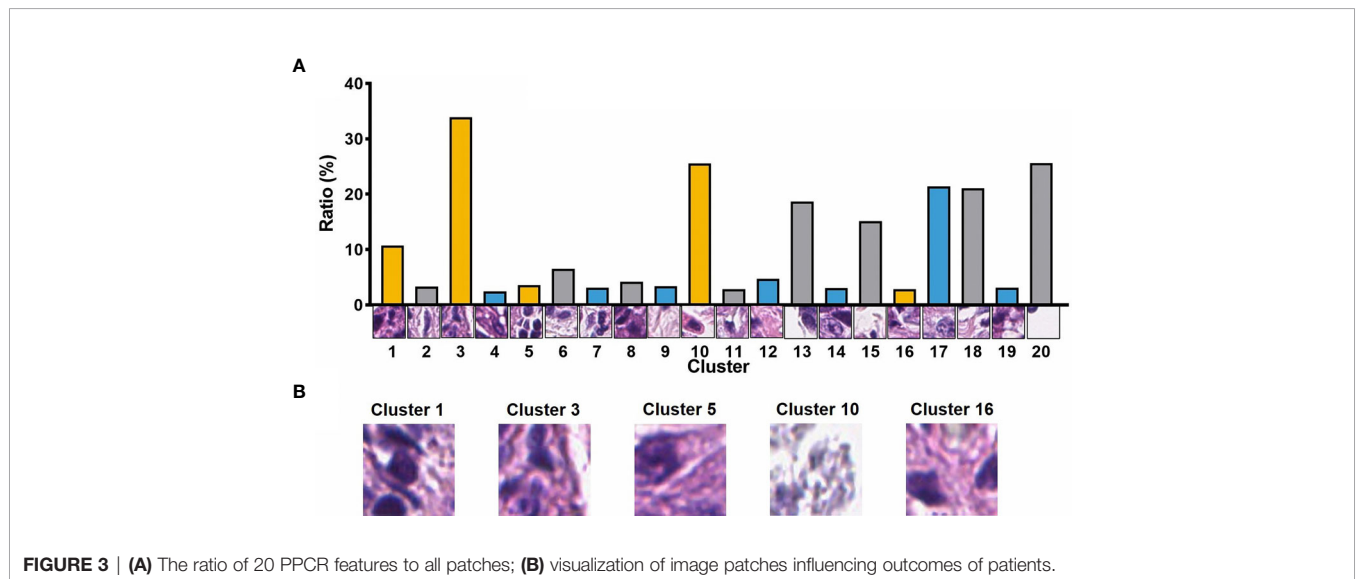

FIGURE 3 | (A) The ratio of 20 PPCR features to all patches; **(B)** visualization of image patches influencing outcomes of patients.

TABLE 3 | Selected features from medical images and clinical information by RFE.

Features	Statistic		
	HR	P value	95% CI
CT features			
original_shape_LeastAxisLength	0.41	0.04	0.02-0.81
original_shape_Maximum2DDiameterColumn	1.07	<0.005	0.46-1.67
original_shape_Maximum3DDiameter	1.47	<0.005	0.75-2.18
original_shape_Sphericity	0.22	0.17	0.09-0.53
original_glszm_GrayLevelNonUniformityNormalized	0.38	0.15	0.14-0.91
original_glszm_LargeAreaEmphasis	0.34	0.01	0.08-0.6
original_glszm_ZoneEntropy	0.85	<0.005	0.31-1.39
original_ngtdm_Contrast	0.3	0.03	0.03-0.57
Pathological features			
Cluster 1	0.01	0.01	0-0.01
Cluster 3	0	0.15	0-0.01
Cluster 5	0.05	0.01	0.01-0.08
Cluster 10	0.01	0.01	0-0.01
Cluster 16	0.04	0.06	0-0.09
Clinical features			
Drinking History	0.32	0.18	0.15-0.79
CRP	0.01	0.01	0-0.02
AGR	1.07	0.05	0.01-2.13
TNM stage	0.14	0.08	0.02-0.3

Stratified Effect of Survival Models

After combining all features into a survival score (SC) as a new feature and classifying the patients into the high-risk and low-risk groups according to the median SC, we found that the stratification improved the CT-pathology-clinical survival model (**Figure 4**, P <0.001). In the CT-pathology-clinical survival model, 1- and 3-year overall survival rates were 38.0% vs. 82.0% and 25.0% vs. 67.0% in the high-risk and low-risk groups, respectively (P <0.001; **Table 4** and **Figure 4**).

DISCUSSION

For analyzing the risks of metastases or recurrence in a cancer patient, the more accurate application of quantitative analysis and correlating it with the knowledge of the individual's prognostic situation can help identify those with a high risk of metastasis and/or recurrence. Radiomic features of the tumor not recognizable by physicians' observations can provide information on both the tumor and its microenvironment. In previous studies, deep learning of radiomics has shown unique value in diagnosis and prognosis in the assessment of liver fibrosis in chronic hepatitis B (24), access to tumor-infiltrating CD8 cells, and responses to anti-PD-1 or anti-PD-L1 immunotherapy (25) that could predict lymph node metastasis in ESCC and colorectal cancer (26, 27). CT texture features may be associated with diagnosis and prognosis in malignant tumors, including lung cancer (7, 9, 28–30). In the quantitative analysis of pathological images, deep learning identified adenocarcinoma or squamous cell carcinoma through the pathological image features of lung cancer that could predict the prognosis of cancer, including colorectal cancer and glioma (11, 31–33). In this study, radiologists sketched the ROIs and the location of ?

A3B2 twb 0.24w?>tumors in esophageal cancer on CT, and pathologists reviewed the HE-stained images and labeled the tumor-enriched areas, then perforated the corresponding wax blocks to create the TMAs. Essentially, most previous studies in pathology that focused on studying a single cell, nuclei, etc., were limited to paraffin tissue sections, which led to changes in the nucleus and imaging results. While studying single cells, it is difficult to fully reflect the true state of the tumor. Therefore, this study aimed to reflect on as many of the features of the tumor on the pathological slides as possible. Therefore, we used whole microarray slide imaging technology to digitize the pathological images, which we could conveniently process by cutting the original picture of every patient into 50×50 size patches from 6000×6000. Then, each patient had a total of 14,400 patches. In addition, we obtained 20 PPCR features after cluster analysis. Eventually, there were five PPCR features reserved. The heatmap showed that Cluster 10 had a color gradient process (**Figure 2**). This study innovatively used cluster analysis to extract pathological features. Using this method of cluster analysis, we processed a large number of pathological images that were convenient and reduced the work of the pathologist; moreover, the features could be automatically identified and sketched by the computer. It is worth mentioning that, as far as we know, this study is the first to use recursive feature elimination (RFE) as a statistical algorithm in the field of medical imaging. The features selected by this algorithm helped us to optimize the model's predictive ability.

As we all know that neoadjuvant therapy (such as chemotherapy, radiotherapy and chemoradiotherapy) is necessary for local advanced ESCC patients to improve survival outcomes (34, 35). For this research, we want to discuss the importance of analyses to diagnostic CT and postoperative histological slides more. Therefore, we excluded the patients who received neoadjuvant therapy to reduce the influence of neoadjuvant therapy on histological images. However, it's also important for patients with advanced stage to receive adjuvant therapy to get a survival benefit. Regrettably, given the flaws of a retrospective study, we unable to obtain more detailed information about adjuvant therapy. Because the part of ESCC patients live far from our cancer center, they may receive adjuvant therapy in the local hospital, but we unable to record this part of the information. Thus, we try to prospectively recruit the ESCC patients to further study the impact of quantitative features to CT and pathological images on prognosis.

We built a clinical survival model based on four features, yielding a C-index adjusted for LOOCV of 0.596, indicating a difference between low-risk and high-risk groups (**Figure 4A**). We found that the CT-clinical model was very effective in distinguishing patient prognoses (**Figure 4B**) and stratified postoperative patients who were prone to recurrence and metastasis. During pathological image processing, we adopted an innovative cluster analysis method, and our pathological-clinical model based on quantitative features was much improved compared to the initial clinical model (P <0.05). The reason that the CT-clinical model was better than the pathology-clinical model might be that a single pathological slice does not fully reflect the biological characteristics
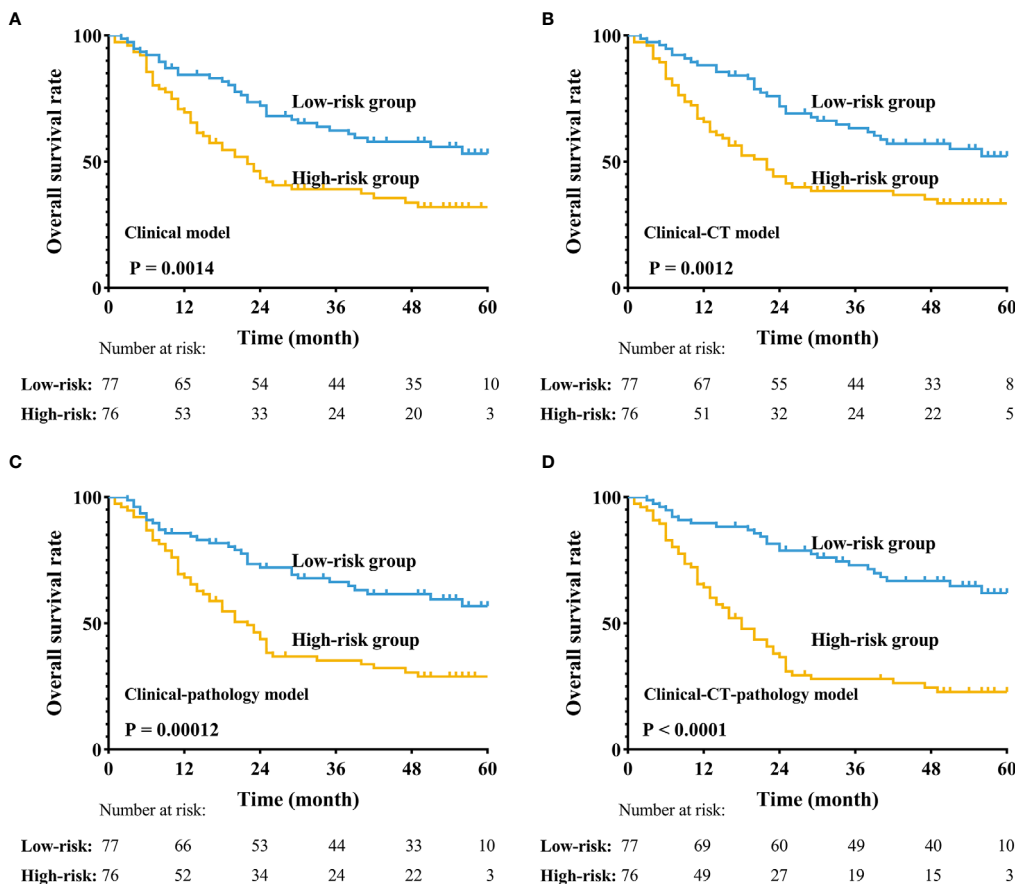
FIGURE 4 | Application of the four different prognostic models to refine the assessment of risk in ESCC. (A–D) Kaplan-Meier survival estimates for a group of patients with ESCC from the whole cohorts and the subgroups predicted to have either a high or a low probability of mortality.

TABLE 4 | Overall survival of patients with ESCC in every model.

Model		Overall Survival (OS)			
		1-year OS (%)	3-year OS (%)	Median survival time (month)	P value
Clinical Model	Low risk	74.0%	58.0%	–	<0.001
	High risk	46.0%	33.0%	22	
CT-clinical model	Low risk	76.0%	57.0%	–	<0.001
	High risk	44.0%	35.0%	22	
Pathology-clinical model	Low risk	73.0%	61.0%	–	0.001
	High risk	47.0%	30.0%	22	
CT-pathology-clinical model	Low risk	82.0%	67.0%	–	<0.001
	High risk	38.0%	25.0%	18	

of the tumor; however, the CT model is based on 3D dimensional features of ESCC. Finally, the CT-pathology-clinical model was significantly better at predicting prognosis than the CT-clinical and pathology-clinical models (P <0.001, **Figure 4**), as it is based on CT, clinical, and pathological data.

Our study combined multidisciplinary data to construct a more comprehensive and more individualized prognostic model in ESCC. Previous studies focused on qualitative observation of images (36–38); by contrast, this study focused on quantitative analysis of CT and pathological images. This was more precise compared to quantitative analysis for creating a survival model, and therefore, more accurate for predicting the prognosis of ESCC patients after surgery.

There are some limitations of this study. First, this was a retrospective study with a small number of cases (153 patients). We verified the stability of our small sample model using a LOOCV; however, it remains necessary to increase the sample size and conduct a prospective study to verify our results. In addition, the distribution of each patient in the staging was unbalanced and represented a skewed distribution. Second, pathologists selected tissue enrichment areas to create the TMA; however, information around the tumor was lost. Furthermore, the results of TMA might lack random effect because of artificial selection; in other words, selecting different tumor areas might have a deviation in experimental results. Third, this study only evaluated the primary tumor in images of CT and pathology; indeed, the status of lymph nodes was also a key indicator to predict the prognosis of ESCC, therefore, further quantitative analysis of medical images with lymph nodes in CT and pathology is needed. Eventually, the results of this study can only reflect the prognosis of patients with postoperative ESCC. Treatment decisions remain the

responsibility of clinicians and need to be based on the individual patient's situation.

In conclusion, quantitative prognostic modeling using a combination of clinical, histopathological, and diagnostic CT data can stratify ESCC patients with surgery alone into high-risk and low-risk groups, improving the prediction of the prognosis of postoperative patients. This comprehensive model included individual information reflecting more accurate conditions of ESCC, assisting precise treatment.

DATA AVAILABILITY STATEMENT

Researchers interested in this study may contact the authors to obtain the clinical data of all 153 patients. We have uploaded the data of this study in the Research Data Deposit of SYSUCC, and its number was RDDB2021001083 (http://www.resea rchda ta.org.cn/).

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by ethics committee of Sun Yat-sen University Cancer Center (Approval No: GZR2019-161). Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

Conception and design: G-WM, YL, JW. Administrative support: PL, HL, L-JZ. Provision of study materials or patients: PL, HL, L-JZ, G-WM. Collection and assembly of data: G-WM, YL, JW, L-LW, WH, XL, J-LW, J-BL, C-YC, JZ, Y-YH. Data analysis and interpretation: G-WM, YL, JW, L-LW, WH, J-LW. Manuscript writing: G-WM, YL, JW, L-LW, WH, J-LW. Final approval of manuscript: G-WM, YL JW, L-LW, WH, XL, J-LW, J-BL, C-YC, JZ, Y-YH. Made of TMA: J-BL, WH, JZ.. Selection of tumor region in CT: G-WM, WH, L-LW, XL, C-YC. All authors contributed to the article and approved the submitted version.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68:394–424 doi: 10.3322/caac.21492
2. Arnold M, Soerjomataram I, Ferlay J, Forman D. Global incidence of oesophageal cancer by histological subtype in 2012. *Gut* (2015) 64:381–7 doi: 10.1136/gutjnl-2014-308124
3. Edgren G, Adami HO, Weiderpass E, Nyrén O. A global assessment of the oesophageal adenocarcinoma epidemic. *Gut* (2013) 62:1406–14 doi: 10.1136/gutjnl-2012-302412
4. Liang H, Fan JH, Qiao YL. Epidemiology, etiology, and prevention of esophageal squamous cell carcinoma in China. *Cancer Biol Med* (2017) 14:33–41 doi: 10.20892/j.issn.2095-3941.2016.0093
5. Abnet CC, Arnold M, Wei WQ. Epidemiology of Esophageal Squamous Cell Carcinoma. *Gastroenterology* (2018) 154:360–73 doi: 10.1053/j.gastro.2017.08.023
6. Du L. Progress research in survival analysis of esophageal squamous cell carcinoma (ESCC) in Chinese population. *Chin J Pract Oncol* (2018) 156:344–8. doi: 10.11904/j.issn.1002-3070.2018.04.012
7. Hawkins SH, Korecki JN, Balagurunathan Y, Gu Y, Kumar V, Basu S, et al. Predicting Outcomes of Nonsmall Cell Lung Cancer Using CT Image Features. *IEEE Access* (2014) 2:1418–26. doi: 10.1109/ACCESS.2014.2373335
8. Song J, Liu Z, Zhong W, Huang Y, Ma Z, Dong D, et al. Non-small cell lung cancer: quantitative phenotypic analysis of CT images as a potential marker of prognosis. *Sci Rep* (2016) 6:38282. doi: 10.1038/srep38282
9. Zhang Y, Oikonomou A, Wong A, Haider MA, Khalvati F. Radiomics-based Prognosis Analysis for Non-Small Cell Lung Cancer. *Sci Rep* (2017) 7:46349. doi: 10.1038/srep46349
10. Mo X, Wu X, Dong D, Guo B, Liang C, Luo X, et al. Prognostic value of the radiomics-based model in progression-free survival of hypopharyngeal cancer treated with chemoradiation. *Eur Radiol* (2019) 10:1–11. doi: 10.1007/s00330-019-06452-w
11. Bychkov D, Linder N, Turkki R, Nordling S, Kovanen PE, Verrill C, et al. Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* (2018) 8:3395. doi: 10.1038/s41598-018-21758-3
12. Huang X, Long L, Wei J, Li Y, Xia Y, Zuo P, et al. Radiomics for diagnosis of dual-phenotype hepatocellular carcinoma using Gd-EOB-DTPA-enhanced MRI and patient prognosis. *J Cancer Res Clin Oncol* (2019) 145:2995–3003. doi: 10.1007/s00432-019-03062-3
13. Xu H, Lv W, Feng H, Du D, Yuan Q, Wang Q, et al. Subregional Radiomics Analysis of PET/CT Imaging with Intratumor Partitioning: Application to Prognosis for Nasopharyngeal Carcinoma. *Mol Imaging Biol* (2019) 10:1–13. doi: 10.1007/s11307-019-01439-x
14. Wei W, Wang K, Liu Z, Tian K, Wang L, Du J, et al. Radiomic signature: A novel magnetic resonance imaging-based prognostic biomarker in patients with skull base chordoma. *Radiother Oncol* (2019) 10:1–8. doi: 10.1016/j.radonc.2019
15. Shen X, Zhao B. Efficacy of PD-1 or PD-L1 inhibitors and PD-L1 expression status in cancer: meta-analysis. *BMJ* (2018) 362:k3529. doi: 10.1136/bmj.k3529
16. Khosravi P, Kazemi E, Imielinski M, Elemento O, Hajirasouliha I. Deep Convolutional Neural Networks Enable Discrimination of Heterogeneous

Digital Pathology Images. *EBioMedicine* (2018) 27:317–28. doi: 10.1016/j.ebiom.2017.12.026

17. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. *Sci Rep* (2017) 7:16878. doi: 10.1038/s41598-017-17204-5

18. Litjens G, Sanchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep* (2016) 6:26286. doi: 10.1038/srep26286

19. Koelzer VH, Sirinukunwattana K, Rittscher J, Mertz KD. Precision immunoprofiling by image analysis and artificial intelligence. *Virchows Arch* (2019) 474:511–22. doi: 10.1007/s00428-018-2485-z

20. Saltz J, Gupta R, Hou L, Kurc T, Singh P, Nguyen V, et al. Spatial Organization and Molecular Correlation of Tumor-Infiltrating Lymphocytes Using Deep Learning on Pathology Images. *Cell Rep* (2018) 23:181–93. doi: 10.1016/j.celrep.2018.03.086

21. Nagpal K, Foote D, Liu Y, Chen PC, Wulczyn E, Tan F, et al. Development and validation of a deep learning algorithm forimproving Gleason scoring of prostate cancer. *NPJ Digital Med* (2019) 2:48. doi: 10.1038/s41746-019-0112-2

22. Wulczyn E, Steiner D, Xu Z, Sadhwani A, Wang H, Flament-Auvigne I, et al. Deep learning-based survival prediction for multiple cancer types using histopathology images *PloS One* (2020) 15(6):e0233678 doi: 10.1371/journal.pone.0233678.

23. Granitto PM, Furlanello C, Biasioli F, Gasperi F. Recursive feature elimination with random forest for PTR-MS analysis of agroindustrial products. *Chemometrics Intelligent Lab Syst* (2006) 83:83–90. doi: 10.1016/j.chemolab.2006.01.007

24. Wang K, Lu X, Zhou H, Gao Y, Zheng J, Tong M, et al. Deep learning Radiomics of shear wave elastography significantly improved diagnostic performance for assessing liver fibrosis in chronic hepatitis B: a prospective multicentre study. *Gut* (2019) 68:729–41. doi: 10.1136/gutjnl-2018-316204

25. Sun R, Limkin EJ, Vakalopoulou M, Dercle L, Champiat S, Han SR, et al. A radiomics approach to assess tumour-infiltrating CD8 cells and response to anti-PD-1 or anti-PD-L1 immunotherapy: an imaging biomarker, retrospective multicohort study. *Lancet Oncol* (2018) 19:1180–91. doi: 10.1016/S1470-2045(18)30413-3

26. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol* (2016) 34:2157–64. doi: 10.1200/JCO.2015.65.9128

27. Tan X, Ma Z, Yan L, Ye W, Liu Z, Liang C. Radiomics nomogram outperforms size criteria in discriminating lymph node metastasis in resectable esophageal squamous cell carcinoma. *Eur Radiol* (2019) 29:392–400. doi: 10.1007/s00330-018-5581-1

28. Ganeshan B, Abaleke S, Young RC, Chatwin CR, Miles KA. Texture analysis of non-small cell lung cancer on unenhanced computed tomography: initial evidence for a relationship with tumour glucose metabolism and stage. *Cancer Imaging* (2010) 10:137–43. doi: 10.1102/1470-7330.2010.0021

29. Wang S, Shi J, Ye Z, Dong D, Yu D, Zhou M, et al. Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning. *Eur Respir J* (2019) 53:1800986. doi: 10.1183/13993003.00986-2018

30. Ardila D, Kiraly AP, Bharadwaj S, Choi B, Reicher JJ, Peng L, et al. End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nat Med* (2019) 25:954–61. doi: 10.1038/s41591-019-0447-x

31. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* (2018) 24:1559–67. doi: 10.1038/s41591-018-0177-5

32. Djuric U, Zadeh G, Aldape K, Diamandis P. Precision histology: how deep learning is poised to revitalize histomorphology for personalized cancer care. *NPJ Precis Oncol* (2017) 1:22. doi: 10.1038/s41698-017-0022-1

33. Mobadersany P, Yousefi S, Amgad M, Gutman DA, Barnholtz-Sloan JS, Velázquez Vega JE, et al. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proc Natl Acad Sci U S A* (2018) 115:E2970–9. doi: 10.1073/pnas.171713911

34. Yang H, Liu H, Chen Y. Neoadjuvant Chemoradiotherapy Followed by SurgeryVersus Surgery Alone for Locally Advanced Squamous Cell Carcinoma of the Esophagus (NEOCRTEC5010): A Phase III Multicenter, Randomized, Open-Label Clinical Trial. *J Clin Oncol* (2018) 36:2796–803 doi: 10.1200/JCO.2018.79.1483

35. National Comprehensive Cancer Network. *Esophageal and Esophagogastric Junction Cancers (Version 4.2020)*. Available at: https://www.nccn.org/professionals/physician_gls/pdf/esophageal.pdf (Accessed 9 March 2021).

36. Jensen HH, Login FH, Park JY, Kwon TH, Nejsum LN, et al. Immunohistochemical evaluation of activated Ras and Rac1 as potential downstream effectors of aquaporin-5 in breast cancer in vivo. *Biochem Biophys Res Commun* (2017) 493:1210–6. doi: 10.1016/j.bbrc.2017.09.125

37. Abdelrahim A, Balmer C, Jones J, Mehanna H, Dunn J. Considerations for a head and neck smoking cessation support programme; A qualitative study of the challenges in quitting smoking after treatment for head and neck cancer. *Eur J Oncol Nurs* (2018) 35:54–61. doi: 10.1016/j.ejon.2018.05.003

38. Di Martino M, Anzidei M, Zaccagna F, Saba L, Bosco S, Rossi M, et al. Qualitative analysis of small (</=2 cm) regenerative nodules, dysplastic nodules and well-differentiated HCCs with gadoxetic acid MRI. *BMC Med Imaging* (2016) 16:62. doi: 10.1186/s12880-016-0165-5

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.