Radiomic analysis for pretreatment prediction of response to neoadjuvant chemotherapy in locally advanced cervical cancer: A multicentre study

Caixia Sun\textsuperscript{a,b,1}, Xin Tian\textsuperscript{c,1}, Zhenyu Liu\textsuperscript{b,d,1}, Weili Li\textsuperscript{c,1}, Pengfei Li\textsuperscript{c}, Jiaming Chen\textsuperscript{c}, Weifeng Zhang\textsuperscript{c}, Ziyu Fang\textsuperscript{c}, Peiyuan Du\textsuperscript{c}, Hui Du\textsuperscript{a}, Ping Liu\textsuperscript{c,2}, Lihui Wang\textsuperscript{a,3}, Chunlin Chen\textsuperscript{c,4}, Jie Tian\textsuperscript{b,d,e,f,3∗∗∗}

\textsuperscript{a} Key Laboratory of Intelligent Medical Image Analysis and Precise Diagnosis of Guizhou Province, School of Computer Science and Technology, Guizhou University, Guiyang, China
\textsuperscript{b} CAS Key Laboratory of Molecular Imaging, Institute of Automation, Chinese Academy of Sciences, Beijing, China
\textsuperscript{c} Department of Gynaecology and Obstetrics, Nanfang Hospital, Southern Medical University, 1838 Guangzhou Avenue North, Guangzhou, China
\textsuperscript{d} Department of Radiology, Southern Medical University, 1838 Guangzhou Avenue North, Guangzhou, China
\textsuperscript{e} Beijing Advanced Innovation Center for Big Data-Based Precision Medicine, School of Medicine, Beihang University, Beijing, China
\textsuperscript{f} School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing, China

\begin{abstract}

Background: We aimed to investigate whether pre-therapeutic radiomic features based on magnetic resonance imaging (MRI) can predict the clinical response to neoadjuvant chemotherapy (NACT) in patients with locally advanced cervical cancer (LACC).

Methods: A total of 275 patients with LACC receiving NACT were enrolled in this study from eight hospitals, and allocated to training and testing sets (2:1 ratio). Three radiomic feature sets were extracted from the intratumoural region of T1-weighted images, intratumoural region of T2-weighted images, and peritumoural region of T2-weighted images before NACT for each patient. With a feature selection strategy, three single sequence radiomic models were constructed, and three additional combined models were constructed by combining the features of different regions or sequences. The performance of all models was assessed using receiver operating characteristic curve.

Findings: The combined model of the intratumoural zone of T1-weighted images, intratumoural zone of T2-weighted images, and peritumoural zone of T2-weighted images achieved an AUC of 0.998 in training set and 0.999 in testing set, which was significantly better (p < 0.05) than the other radiomic models. Moreover, no significant variation in performance was found if different training sets were used.

Interpretation: This study demonstrated that MRI-based radiomic features hold potential in the pretreatment prediction of response to NACT in LACC, which could be used to identify rightful patients for receiving NACT avoiding unnecessary treatment.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

\end{abstract}

1. Introduction

Cervical cancer is one of the most frequent malignant tumours in women\cite{1}. Standard therapeutic strategies for locally advanced cervical cancer (LACC) include radical hysterectomy (RH) or concurrent chemoradiation\cite{2}. Preoperative neoadjuvant chemotherapy (NACT) is being increasingly investigated as alternative treatment strategy for locally advanced disease due to its ability to reduce the tumour volume and render unresectable tumours operable\cite{3–5}. Chemotherapeutic responders could exhibit clinical survival benefit over patients undergone primary surgery\cite{6,7}. Conversely, chemotherapeutic non-responders experience unnecessary chemotherapy-related toxicities and disease
Radiomics, as a newly-proposed method, aims to evaluate tumour heterogeneity by extracting high throughput features from medical images that reflect the underlying pathophysiology [18–20]. These radiomic features capture distinct phenotypic differences and may have prognostic and predictive value across different diseases [21–24]. Emerging studies have demonstrated the excellent performance of radiomic models for predicting the pathological complete response (pCR) to NACT in rectal cancer [25,26]. Prediction of the pCR to NACT in breast cancer was successfully performed by radiomic analysis using pretreatment breast MRI in several studies [27,28]. In the field of cervical cancer, El Naqa et al. [29] reported the good predictive power of radiomic features in predicting individual chemoradiotherapy outcomes based on single centre small patient cohorts. However, most of the established predictive models were trained with a single centre dataset may have limited generalization capabilities. Until recently, Lucia et al. [30,31] further verified the prognostic role of pretreatment MRI-based radiomics in patients with LACC treated with definitive chemoradiation using external multicentre patient datasets. These studies provided the likelihood of employing radiomics in patients with cervical tumour for prediction of the response to NACT.

Based on the successful application of radiomics in other research areas, we hypothesized that pretreatment radiomic features would be effective in the prediction of the response to NACT in patient with LACC. Thus, in the present study, we aimed to construct a model for predicting the response to NACT by exploring multi-centric pretreatment MRI-based radiomic analysis, and validate the model with a multicentre dataset.

2. Materials and methods

2.1. Patients

This study design was approved by an institutional review board and waived the informed consent requirement. The patients of this retrospective multicentre study were enrolled in from eight centres between January 2009 and June 2018. In this study, 275 consecutive patients with histologically confirmed LACC who underwent NACT and had pretreatment MRI data were included. The recruitment of patients is described in the supplemental material and Fig. S1. In this study, patients were divided into training and testing sets according to the hospital they were recruited from. The training set consisted of 183 patients from five hospitals, including the first affiliated hospital of Zhengzhou university, second affiliated hospital of zhengzhou university, nanfang hospital, yuncheng central hospital, affiliated hospital of qingdao university. The independent-testing set includes 92 patients from the remaining three hospitals, including the hanen provincial people's hospital, army military medical university xinqiao hospital, yantai sulphur top hospital.

All of the included patients received cisplatin-based (70–75 mg/m2) chemotherapy intra-arterially or intravenously. The tumour volume was evaluated by a gynecological examination and pelvic MRI scan before each chemotherapy cycle and three weeks after NACT. Besides, the extent of change in tumour volume needs to be confirmed by pathological examination. When the patient’s tumour shrinks to meet the surgical conditions, the patient would underwent Type III RH after one cycle of chemotherapy. The surgery was performed within 21 days after the end of the final cycle. In contrast, the patients who did not meet the surgical requirements received 1–2 cycles of chemotherapy every 21 days. In addition, radiotherapy was performed for patients with progressive diseases or those who experienced adverse side effects due to the toxicity of anti-tumour agents during or after NACT. In this study, patients from different centres underwent distinct courses of
NACT. Among these patients, 131 patients underwent one cycle of NACT, 126 patients underwent two cycles, and 18 patients had three cycles. The short-term response was assessed by change in tumour volume according to the Response Evaluation Criteria in Solid Tumours (RECIST v.1.1) [32]. We defined responsive/wess according to the various treatment courses of patients in this study as follows: patients with a complete response or partial response within two cycles of chemotherapy were considered as "responders" (chemo-sensitive), while those still with progressive disease or stable disease after two cycles of treatment were considered as "non-responders" (chemo-insensitive) regardless of the response after the final chemotherapy cycles. Thus, there were 204 responders and 71 non-responders among the 275 eligible patients.

2.2. Image acquisition and tumour masking

Pretreatment MRI examinations were performed within one week before the first cycle of NACT. The MRIs of training set combined five hospitals were performed with 3.0-T (Skyra, SIEMENS, Germany), 1.5-T (Siemens Excite, Germany), 1.5-T (Symphony Tim, SIEMENS, Germany) and 1.5-T (Signa HDxt, GE, USA). The MRIs of testing set combined three hospitals were performed with 3.0-T (TrioTim, SIEMENS, Germany), 3.0-T (Signa HDxt, GE, USA) and 3.0-T (Signa Excite, GE, USA). The MRI protocols of each hospital included axial TSE T2-weighted fat-suppressed sequence and T1-weighted sequence but have some different. The comprehensive parameters of MR image acquisition of each hospital are provided in Table S1.

The pretreatment MRI sequences were collated for tumour segmentation and feature generation. The regions of interest (ROIs) of images were manually delineated on each slice (using ITK-SNAP software: www.itksnap.org) containing the tumour mass on T1-weighted imaging (T1WI) and fat suppression on T2-weighted imaging (T2WI). The T2WI series delineated the peritumoural area in addition to the intratumoural region. The inter- and intra-observer reproducibility for ROI-based feature extraction were analysed in a blind fashion by two radiologists with 7 (reader 1) and 10 (reader 2) years of experience. More details regarding reproducibility are presented in the supplementary material.

2.3. Radiomic feature extraction and selection

Radiomic feature extraction was conducted using the in-house toolbox developed by MATLAB 2016b (Mathworks, Natick, MA, USA). Before the feature extraction, each slice of MR images was normalized with z-score so that get a standard normal distribution of image intensities. Then, six hundred and forty seven predefined features were extracted from each ROI of the MR images: 17 first order statistic features, eight shape features, 54 textural features, and 568 wavelet features (Table S2). The 647 features were extracted from the intratumoural region on T1WI, intratumoural region on T2WI, and peritumoural region on T2WI. Finally, a total of 1941 radiomic features were extracted from each patient. The extracted features are reproducible and match the benchmarks of IBSI [33,34]. Each feature of each patient in training or testing sets was normalized with z-score in order to remove the effect of different MRI scanners [35].

Introducing a large number of features to the prediction model may cause over-fitting, thereby reducing the performance of the model on another dataset. To reduce the dimensionality of features, we adopted recursive feature elimination based on a support vector machine (SVM-RFE) to select top features in three feature sets [36]. Previous studies [37,38] have demonstrated that SVM-RFE was excellent on radiomic analysis for feature selection, which aims to select the features that best fit the current classification task. The SVM-RFE is used to determine the optimal subset of features through calculating the classified accuracy with different combinations of features by SVM.

2.4. Single sequence model construction and validation

To constructed the radiomic models, the previously selected features were used as inputs to random forest (RF) models. These models were constructed based on the training set, and a grid search with three-fold cross-validation was applied to set the hyperparameters of the RF models. After the models were constructed, their ROC curves were plotted to assess and compare the quantitative performance of different models.

Three single MRI sequence models were constructed (T1WI with the intratumoural zone, T2WI with the intratumoural zone, T2WI with the peritumoural zone) within the training set, and the performance of the models was validated with testing set.

2.5. Combined model construction and validation

Three different combined radiomic models were generated by combining the features of different sequences or tumour regions, including combination of intratumoural T1WI and intratumoural T2WI, T2WI sequence with both intratumoural and peritumoural regions, and T1WI and T2WI sequences with all regions (multi-sequence). It should be emphasized that the construction process of the combined models was same as that for single sequence models. The pipeline of the entire experiment is presented in Fig. 1.

In addition, to assess the robustness of the models, three additional different training sets were composed of patients from three random combinations of hospitals, and the patients from remaining hospitals were assigned to testing sets. The hospital combination between each group of training and testing sets were different, the combinations also differ from original training and testing sets, which to verify whether the predictive power of the model depends on the patient composition of the training set. As indicated in Table S3, all training and testing sets comprised a similar number of patients. The training and testing process of the three patient groups were used the same features selection and model construction strategy as in the previous experiment. Then the robustness of models for NACT response prediction was evaluated by the ROC curve.

2.6. Clinical factor and radiomic model

A clinical model was constructed based on the subset of patients whose clinical information (age, FIGO stage, and gross type) is available. The process of clinical model construction was consistent with radiomic model. Then, as a comparison, evaluating the predictive result of the patients in radiomic model.

2.7. Statistical analysis

Descriptive statistics of continuous variables are presented as the mean ± standard deviation, and the Student's t-test was applied to compare the difference between groups. The p value of the χ2 test was used to illustrate differences between qualitative variables. Generally, p values were two sided, and p < .05 was considered statistically significant. Comparison of qualitative variables between groups was conducted with a one-way analysis of variance. The cut-off values of models were determined according to the Youden index [40] of training sets to quantize the discrimination ability of models. Quantitative comparison of AUCs was made with the Delong test [41]. The predictive accuracy of models was evaluated by balanced accuracy [42] due to the imbalanced ratio between responders and non-responders. All experiments and statistical analyses were implemented with python 2.7.

2.8. Data statement

The experimental data are not available for public access due to patient privacy concerns but can be obtained from the corresponding
Table 1
Clinical characteristics of patients in the training and testing sets.

| Characteristics | Training set   | Testing set   | p       | Testing set   | p       |
|-----------------|----------------|---------------|---------|---------------|---------|
|                 | Responders     | Non-responders|         | Responders    |         |
| Age (years, mean ± SD) | 47.9 ± 8.1     | 46.1 ± 9.2    | 0.227   | 50.9 ± 7.7    | <0.001* |
| FIGO stage      |                |               |         |               |         |
| IB2             | 57 (38.3%)     | 11 (32.4%)    |         | 7 (12.7%)     |         |
| IIA-IIA2        | 62 (41.6%)     | 6 (17.6%)     |         | 16 (29.1%)    |         |
| IIIB-IIIB       | 30 (20.1%)     | 17 (50.6%)    |         | 32 (58.2%)    |         |

Note: The chi-square test was used to compare the difference in categorical variables (FIGO stage), while a Student’s t-test was used to compare the difference in age.

* P < .05. SD, standard deviation.
zone of T2WI, and NGTDM_complexity feature in the peritumoural zone of T2WI. It can be observed that all features of the responders have higher expression than those of non-responders.

3.3. Single sequence model construction and validation

Based on the selected features, the radiomic models for each sequence with different tumour regions were constructed using the RF method, including the T1WI sequence with the intratumoural region, a T2WI sequence with the intratumoural region, and a T2WI sequence with the peritumoural region.

The ROC curves of the above-mentioned models in the training and testing sets are presented in Fig. 3a–b. Using the radiomic features extracted from the intratumoural region, the AUCs for single-sequence (T1WI or T2WI) models were >0.96 in the training set and >0.94 in the testing set. However, using the radiomic features extracted from the peritumoural region, the AUC for the single sequence (T2WI) model could reach 0.975 in the training set and 0.98 in the testing set. The specificity of three models was show good performance in training and testing sets. Different from specificity, the sensitivity of these models was >84% in the training set but <77% in the testing set. To further evaluate the performance of different models, the quantitative metrics are provided in Table 2.

3.4. Combined model construction and validation

The three combined models were then constructed by combining features of different sequences or regions, including intratumoural combination of T1WI and T2WI, T2WI with both peritumoural and intratumoural regions, and T1WI and T2WI with all tumour regions (multi-sequence). The performance of the three combined models are illustrated in Fig. 3c–d. From Table 3, we observed that, using the combination of radiomic features extracted from intratumoural T1WI and intratumoural T2WI, the AUC for the combined model was 0.967 and 0.980 in the training and testing sets, respectively. Meanwhile, another model, combining both tumour regions of T2WI, yielded an AUC of 0.991 in the training set and 0.994 in the testing set. Considering the significance of multi-sequence images in feature expression and compensation, the response prediction using all radiomic features extracted from multi-sequences images was evaluated, the AUC of which was further increased to 0.998 in the training set and 0.999 in the testing set. Specifically, the combined model of intratumoural regions (T1WI and T2WI) and combined model of both regions (T2WI) yielded good sensitivity in the training set but exhibited an unsatisfactory performance in the testing set. However, the multi-sequence model not only exhibited excellent sensitivity and specificity in the training set, but also performed satisfactorily in the testing set.

![Feature expression maps for top radiomic features.](image)

- **(a) and (d)** are Coif3_glszm_low grey-level zone emphasis features of intratumoural T1WI of patients.
- **(b) and (e)** are Coif1_ngtdm_busyness features of intratumoural T2WI of patients.
- **(c) and (f)** are Coif3_ngtdm_complexity features of peritumoural T2WI of patients.

T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.
3.5. Model comparison

The performance of the multi-sequence model was significantly increased \((p < .05)\) compared to most single sequence models in both sets (Fig. 4). In the comparison of combined models in the training set, the AUC of the multi-sequence model was higher \((p < .05)\) than that of the other combined models. Further, in the testing set, there were no significant differences in the AUCs among the three combined models.

Furthermore, for each group of training and testing sets consisting of patients from different combination of hospitals, the three ROC curves (Fig. 5) of the combined models were robust. The performances of the models in the three combination groups have no significantly difference (Table S8). Similar to the previous experimental results, the sensitivity and specificity of the multi-sequence model were more stable than those of the other combined models in both sets. This indicated that regardless of the source of the training data, the multi-sequence model exhibited robust and persistent performance.

3.6. Clinical factor and radiomic model

The clinical factors were identified as independent predictors for the response to NACT, including age, FIGO stage, gross type. Finally, only 169 and 63 patients have complete clinical information in training and testing sets. As shown in Fig. 6, the clinical model yielded AUCs of

Table 2

| Cohort        | Model      | ACC (%) | AUC     | SEN (%) | SPE (%) | Cut-off |
|---------------|------------|---------|---------|---------|---------|---------|
| Training set  | Intratumoural T1WI | 91.7 (85.6–96.9) | 0.964 (0.931–0.996) | 95.9 (91.8–98.1) | 91.2 (86.3–97.6) | 0.643 |
|               | Intratumoural T2WI | 89.5 (84.9–91.2) | 0.966 (0.943–0.989) | 84.6 (75.2–88.0) | 97.1 (90.1–100) | 0.660 |
|               | Peritumoural T2WI | **93.7** (88.9–97.7) | **0.975** (0.956–0.993) | 93.3 (89.2–97.3) | 94.1 (85.2–98.6) | 0.556 |
| Training set  | Intratumoural T1WI | 77.3 (70.9–83.7) | 0.960 (0.942–0.997) | 49.1 (41.9–67.3) | 100 (98.9–100) | 0.643 |
|               | Intratumoural T2WI | 69.1 (62.3–75.8) | 0.945 (0.904–0.986) | 41.8 (33.8–61.4) | 100 (99.1–100) | 0.669 |
|               | Peritumoural T2WI | **87.3** (80.9–93.1) | **0.979** (0.959–0.999) | **76.4** (61.8–86.2) | **100** (97.5–100) | 0.556 |

Statistical quantifications are reported with 95% confidence intervals. T1WI, T1-weighted imaging; T2WI, T2-weighted imaging; ACC, balanced accuracy; AUC, area under receiver operating characteristic curve; SEN, sensitivity; SPE, specificity.
0.666 and 0.608 in training and testing sets. The predictive result of clinical model was significantly lower 30% than radiomic models.

4. Discussion

In this multicentre study, we investigated radiomic analysis based on pretreatment MRI scans to predict the clinical response to NACT in patients with LACC. A multi-sequence model combining the intratumoural and peritumoural regions of two radiomic sequences demonstrated the best predictive power in both training (AUC: 0.998) and testing (AUC: 0.999) sets.

In the present study, ten features were selected from the intratumoural region on T1WI, two features were selected from the intratumoural region on T2WI, and four features were selected from the peritumoural region on T2WI. Among these features, most were textural features of images, which allow multi-view description of the tumour phenotype and cannot be easily identified by humans. From the perspective of view of features, the result was similar to those of previous studies of other cancers [43,44], the features of responders presented higher textural complexity or heterogeneity than those of non-responders. After the single sequence models were constructed, all models demonstrated relatively satisfactory predictive performance. Strikingly, although the peritumoural T2WI model did not carry the highest number of features, it still exhibited higher predictive power than other single sequence models. In previous studies, Braman et al. [43] confirmed the valuable of peritumoural textural radiomic features in enhancing pCR prediction by quantitatively characterizing the presence of tumour infiltrating lymphocytes (TILs) and surrounding stromal abnormalities or responses. The study demonstrated that the peritumoural radiomics profile correspondingly captured the histological heterogeneity of TILs and micro-vascular invasion between non-pCR and pCR. Similarly, many studies [45] have proposed that the pretreatment host immune states of advanced cervical cancer could affect the treatment response to NACT. Moreover, Yun et al. [46] recently discovered a significant correlation between low peritumoural Foxp3+ infiltrating T cells and the clinical efficacy of NACT in patients with LACC. Therefore, peritumoural immunological heterogeneity between chemotherapeutic non-responders and responders was characterized by the best-performing radiomic features.

In addition, we also constructed combined models to study the impact of the combination of different regions or sequences on predictive power. The performance of the combined model of intratumoural regions was highly similar to that of the T1WI model, with no significant improvement. Surprisingly, the T2WI with intratumoural and peritumoural regions model exhibited significantly enhanced predictive performance compared to the single sequence model, with high AUCs in the training and testing sets. However, the performance of these models in terms of sensitivity was not satisfactory. They had excellent sensitivity in training set but not in testing set. Therefore, the multi-sequence model was generated, which combined all sequences and their tumour regions. The multi-sequence model exhibited the best predictive ability, which was significant higher (p < 0.05) than all other models. The accuracy, AUC, sensitivity and specificity of the multi-sequence model offered outstanding predictive power and generalization in both the training and testing sets. This indicates that the multi-sequence model can not only accurately predict the response to NACT, but also perfectly identify non-responders. This will help to protect these non-responders from NACT before treatment and avoid a delayed treatment process. By incorporating radiomic features reflecting peritumoural and tumoural heterogeneity, the multi-sequence model offered the most comprehensive and thorough radiomic assessment for predicting the clinical response to NACT and yielded excellent performance. Our results

![Fig. 4. Delong test between different models. P-value of Delong test between any two models. T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.](image)

**Table 3**

Performance of combined models.

| Cohort      | Model                                      | ACC (%) [95% CI] | AUC [95% CI] | SEN (%) [95% CI] | SPE (%) [95% CI] | Cut-off |
|-------------|--------------------------------------------|------------------|--------------|------------------|------------------|---------|
| Training    | Intratumoural T1WI + Intratumoural T2WI   | 93.4 (88.9–96.6) | 0.967 (0.929–1) | 93.3 (89.2–97.2) | 94.1 (84.6–100)  | 0.658   |
|             | Intratumoural T2WI + peritumoural T2WI    | 93.4 (88.9–96.6) | 0.991 (0.981–1) | 91.9 (85.7–95.5) | 100 (94.3–100)   | 0.670   |
|             | Multi-sequence                             | **97.8 (94.5–99.4)** | **0.998 (0.994–1)** | **97.3 (94.2–99.4)** | **100 (90.6–100)** | **0.594** |
| Testing     | Intratumoural T1WI + Intratumoural T2WI   | 67.4 (56.8–76.8) | 0.980 (0.959–1) | 45.5 (32.2–58.5)  | 100 (93.6–100)   | 0.658   |
|             | Intratumoural T2WI + peritumoural T2WI    | 80.4 (70.9–87.9) | 0.994 (0.986–1) | 67.3 (50.6–81.3)  | 100 (98.1–100)   | 0.670   |
|             | Multi-sequence                             | **90.2 (82.2–95.4)** | **0.999 (0.997–1)** | **83.6 (76.2–93.3)** | **100 (97.5–100)** | **0.594** |

Statistical quantifications are reported with 95% confidence intervals. T1WI, T1-weighted imaging; T2WI, T2-weighted imaging; ACC, balanced accuracy; AUC, area under receiver operating characteristic curve; SEN, sensitivity; SPE, specificity.
provided additional evidence supporting the use of MRI in predicting the clinical response to NACT. We investigated the diagnostic robustness for response prediction of the combined models, and three various training sets constructed by different combinations of centres were applied. As indicated in Table S8, the multi-sequence model performed strong robustness in different experimental groups. The predictive accuracy of the multi-sequence model for non-responders was higher and more stable than that of other models. By comparing the information of features, some features (Table S9) were repeatedly selected in different experimental groups, such as glszm_GLV extracted from

Fig. 5. Comparison of ROC curves between different group of training and testing sets for combined models. (a) and (b) are ROC curves of combined model of intratumoural region on T1WI and intratumoural region on T2WI in training and testing sets. (c) and (d) are ROC curves of combined model of peritumoural and intratumoural zones on T2WI in training and testing sets. (e) and (f) are ROC curves of multi-sequence model in training and testing sets. ROC receiver operating characteristic; AUC area under receiver operating characteristic curve; T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.

Fig. 6. Comparison of ROC curves between clinical and radiomic models. ROC receiver operating characteristic; AUC area under receiver operating characteristic curve; T1WI, T1-weighted imaging; T2WI, T2-weighted imaging.
T1WI, ntgdm.busyness extracted from intratumoural zone of T2WI, glcm_correlation extracted from intratumoural zone of T2WI, and ntgdm_complexity extracted from peritumoural of T2WI. This shows that the strong predictive power and robustness of these radiomic features in prediction of response to neoadjuvant chemotherapy in locally advanced cervical cancer. Furthermore, a clinical model was constructed based on the subset of patients whose clinical information is available. From the experimental result, the clinical factor cannot accurately predict the response to NACT.

Currently, the treatment response in cervical cancer is determined by the shrinkage of tumour size while MRI is the most effective imaging modality to assess tumour response to chemotherapy. As morphological change is not evident until completion of certain cycles of chemotherapy, chemotherapeutic non-responders cannot be earlier ruled out, resulting in increased toxicity, morbidity and worse prognosis. Some proposed imaging biomarkers from functional MRI and cellular markers such as P-Glycoprotein and proliferating cell nuclear antigen [47], squamous cell carcinoma antigen [14], and cyclooxygenase (COX)-2 ratio [48] are either limited by post-therapeutically basis or lack of validation. So far, clinical use of NACT is hampered by the lack of effective and validated pre-therapy predictors to identify suitable chemotherapeutic responders. Recently, results of phase III randomized trial [49] revealed comparable efficacy between NACT-RH and chemoradiation emphasizing the importance of remaining quality of life in treatment planning. In selected chemo-sensitive LACC patients, NACT-RH are likely to be an optimal treatment plan avoiding function loss of vaginal, ovary caused by radical radiotherapy. Therefore, a non-invasive MRI-based pre-treatment radiomics model that allows for accurate and rapid prediction of individual therapy response presented in the study would have immense value in clinical practice. Besides, several prior researches [50,51] have investigated the predictive potential of radiomic features by combining FDG PET and MRI in lung cancer and head-and-neck cancer. Meanwhile, some metabolic radiomics features that characterizing intratumoural metabolic heterogeneity has shown to be valuable for assessing locoregional recurrences and distant metastasis after definitive chemoradiotherapy in LACC [31,52]. The further utility of radiomics based on FDG PET and MRI in the pre-therapeutically prediction of response to NACT in LACC may help to discover radiomic predictors of chemotherapeutic response that potential possesses unique prognostic information.

Our study still has some limitations despite the encouraging results. Firstly, due to the MR images being retrospectively collected from various centres with different scanners and acquisition parameters, we finally adopted a limited population size with all required sequences for model construction. Secondly, we simplified patients’ histocytes into squamous cell carcinomas only for the purposes the investigation. Therefore, the radiomic model might not be generalizable to adenocarcinoma or neuroendocrine carcinoma. Further studies with larger independent external datasets from other centres are warranted to confirm our proposed radiomic models. If possible, prospective studies are needed for further validation.

5. Conclusion

The clinical utility of neoadjuvant therapy in the treatment of advanced cervical cancer is hampered due to lack of effective predictors of response prior to the initiation of therapy. Such a predictive tool could allow for prudent selection of suitable patients for NACT while identifying non-responders for more personalized care avoiding ineffective therapies. For this purpose, we successfully combined peritumoural and intratumoural radiomics from highly heterogeneous pretreatment MRI scans to predict the clinical response to NACT with superior accuracy. The good predictive robustness of the proposed model among various centres with consistent accuracy indicated its potential for widespread adoption in clinical decision-making for NACT in patients with LACC. Our findings suggested that the radiomic model is sufficient to serve as an effective tool to stratify LACC patients to allow for more appropriate initial treatment, showing promise in improving the management of LACC.

Funding sources

This work was supported by the National Key Research and Development Plan of China [Grant No. 2017YFA0205200]; the National Natural Science Foundation of China [Grant No. 81772012, 81272901, 81527805, and 66161010]; the Nature Science Foundation of Guizhou province [Grant No. 20152044]; the Chinese Academy of Sciences [Grant No. GJSTD20170004, KDB201302001, and QYZD-JSC005]; the Beijing Natural Science Foundation [Grant No. 7182109]; the Youth Innovation Promotion Association CAS [Grant No. 2019136]; The National Natural Science Foundation of Guangdong [Grant No. 2015A030311024]; the Health and Medical Cooperation Innovation Special Program of Guangzhou Municipal Science and Technology [No. 201508020264]; the National Key Technology Program of the Ministry of Science and Technology [863 program, Grant No. 2014BA010503]; the Medical Scientific Research Foundation of Guangdong Province of China [Grant No. A2015063]. The funders had no role in study design, data collection, data analysis, interpretation, writing of the report.

Declaration of Competing Interests

The authors declare no potential conflicts of interest.

Author contributions

Conception and design: Chunlin Chen, Jie Tian. Collection and assembly of data: Xin Tian, Weili Li, Pengfei Li, Jianming Chen, Weifeng Zhang, Ziyu Fang, Peiyuan Du, Hui Duan. Development of methodology: Chunlin Chen, Lihui Wang, Ping Liu, Weili Li, Zhenyu Liu, Caixia Sun, Xin Tian. Data analysis and interpretation: Caixia Sun, Xin Tian, Zhenyu Liu, Weili Li, Ping Liu, Lihui Wang, Chunlin Chen, Jie Tian. Manuscript writing: All authors. Final approval of manuscript: All authors.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ebiom.2019.07.049.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019;69:7–34. https://doi.org/10.3322/caac.21551.
[2] Bhattacharya N, Aoki D, Sharma DN, Sankaranarayanan R. Cancer of the cervix uteri. Int J Gynecol Obstet 2018;143:22–36. https://doi.org/10.1016/j.ijgo.201611.
[3] DeSouza NM, Souther WP, Rustin G, Mahon MM, Jones B, Dina R, et al. Use of neoadjuvant chemotherapy prior to radical hysterectomy in cervical cancer: monitoring tumour shrinkage and molecular profile on magnetic resonance and assessment of 3-year outcome. Br J Cancer 2004;90:2326–31. https://doi.org/10.1038/sj.bjc.6601870.
[4] Yang Z, Chen D, Zhang J, Yao D, Gao K, Wang H, et al. The efficacy and safety of neoadjuvant chemotherapy in the treatment of locally advanced cervical cancer: a randomized multicenter study. Gynecol Oncol 2016;141:231–9. https://doi.org/10.1016/j.ygo.2015.06.027.
[5] Angioli R, Piotto F, Monteria R, Aloisi A, Luvero D, Capriglione S, et al. Neoadjuvant chemotherapy plus radical surgery followed by chemotherapy in locally advanced cervical cancer. Gynecol Oncol 2012;127:290–6. https://doi.org/10.1016/j.ygo.2012.07.104.
[6] Hu T, Li S, Chen Y, Shen J, Li X, Huang K, et al. Matched-case comparison of neoadjuvant chemotherapy in patients with FIGO stage Ib1-IIb cervical cancer to establish selection criteria. Eur J Cancer 2012;48:2353–60. https://doi.org/10.1016/j.ejca.2012.03.015.
[7] Scambia G, Benedetti Panici P, Forni E, Amoroso M, Salerno G, Ferrandina G, et al. Squamous cell carcinoma antigen: prognostic significance and role in the monitoring of neoadjuvant chemotherapy response in cervical cancer. J Clin Oncol 1994;12:2309–16. https://doi.org/10.1200/JCO.1994.12.11.2309.
[8] Chen H, Liang C, Zhang L, Huang S, Wu X. Clinical efficacy of modified preoperative neoadjuvant chemotherapy in the treatment of locally advanced (stage IB2 to IB) breast cancer: a multicenter study. Gynecol Oncol 2018;119:308–15. https://doi.org/10.1016/j.ygyno.2018.05.026.

[9] Katsumata N, Yoshikawa H, Kobayashi H, Saito T, Kuzuya K, Nakanishi T, et al. Phase III randomized controlled trial of neoadjuvant chemotherapy plus radical surgery vs radical surgery alone in stages IB2 and IB cervical cancer: a Japan clinical oncology group study. JCOG (JC02). Br J Cancer 2013;108:1957–63. https://doi.org/10.1038/ijc.2013.179.

[10] Harry VN, Semple SJ, Gilbert FJ, Parkin DE. Diffusion-weighted magnetic resonance imaging for the early detection of response to chemotherapy in cervical cancer. Gynecol Oncol 2008;111:213–20. https://doi.org/10.1016/j.ygyno.2008.07.048.

[11] Tanderpak K, Foldal LU, Sturdza A, Haie-Meder C, Mazeron R, van Limbergen E, et al. Effect of tumor dose, volume and overall treatment time on local control after radiotherapy including MRI guided brachytherapy of locally advanced cervical cancer. Radioter Oncol 2016;120:441–6. https://doi.org/10.1016/j.radonc.2016.05.014.

[12] Zhang W, Chen C, Liu P, Li W, Hao M, Zhao W, et al. Impact of pelvic MRI in routine clinical practice on staging of IB1-IB2 cervical cancer. Cancer Manage Res 2019;11:3683–9. https://doi.org/10.2147/CMAJ.S197496.

[13] Zhang W, Chen C, Liu P, Li W, Hao M, Zhao W, et al. Staging early cervical cancer in China: data from a multicenter collaborative. Int J Gynecol Cancer 2019;29:869–73. https://doi.org/10.1136/ijgc-2019-003536.

[14] Yin M, Hou Y, Zhang T, Cui C, Zhou X, Sun F, et al. Evaluation of chemotherapy response with serum squamous cell carcinoma antigen level in cervical cancer patients: a prospective cohort study. PloS One 2013;8. https://doi.org/10.1371/journal.pone.0064992.

[15] Fu C, Bani D, Liu F, Feng X, Du W, Wang X. The value of diffusion-weighted magnetic resonance imaging in assessing the response of locally advanced cervical cancer to neoadjuvant chemotherapy. Int J Gynecol Cancer 2012;22:1037–43. https://doi.org/10.1097/IGC.0b013e3182275367.

[16] Himoto Y, Fujimoto K, Kido A, Matsumura N, Baba T, Daido S, et al. Assessment of the treatment effects among ART-naïve HIV patients: a multivariate pattern analysis of DTI data. Theranostics 2019;9:1303–13. https://doi.org/10.7150/thno.30309.

[17] Zhang Y, Xu X, Tian Q, Li B, Wu Y, Zhang Y, et al. Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging. J Magn Reson Imaging 2017;46:1281–8. https://doi.org/10.1002/jmri.26327.

[18] Tian Y, Liu Z, Tang Z, Li M, Lou X, Dong E, et al. Radiomics analysis of DTI data to assess vision outcome after intravenous methylprednisolone therapy in Neuromyelitis optica neuritis. J Magn Reson Imaging 2018;49:1365–70. https://doi.org/10.1002/jmri.26326.

[19] Lucia F, Visvikis D, Miranda O, Desseroit MC, Robin P, Pradier O, et al. PO-0805: predictive Radiomic models for data from independent institutions using normalized features. IEEE Trans Radiat Plasma Med Sci 2019;3:310–5. https://doi.org/10.1109/TPRMS.2019.2893860.

[20] Huang ML, Hung YH, Lee WM, Li RK, Jiang SY. SVM-RFE based feature selection and taguchi parameters optimization for multiclass SVM classifier. Sci World J 2014, 2014. https://doi.org/10.1155/2014/759624.

[21] Zhang X, Xu X, Tian Q, Li B, Wu Y, Zhang Y, et al. Radiomics assessment of bladder cancer grade using texture features from diffusion-weighted imaging. J Magn Reson Imaging 2017;46:1281–8. https://doi.org/10.1002/jmri.26327.

[22] Tian Y, Liu Z, Tang Z, Li M, Lou X, Dong E, et al. Creating robust predictive Radiomic models for data from independent institutions using normalization. IEEE Trans Radiat Plasma Med Sci 2019;3:310–5. https://doi.org/10.1109/TPRMS.2019.2893860.

[23] Youdel WD. Index for rating diagnostic tests. Cancer 1950;32:5–2.

[24] Dong D, Elizabeth R, July M. Delong DLC–P: comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–45.

[25] Brodersen KH, Ong CS, Stephan KE, Buehmann JM. The balanced accuracy and its posterior distribution. Proc - Conf Int Conf Pattern Recogn 2010;31:312–4. https://doi.org/10.1109/ICPR.2010.764.

[26] Braman NM, Etesami M, Prasanna P, Dubchak C, Gilmore H, Tiwari P, et al. Intratumoral and peritumoral radiomics for the prediction of histopathological complete response to neoadjuvant chemotherapy based on breast DCE- MRI. Breast Cancer Res 2017;19:1–14. https://doi.org/10.1186/s13058-017-0846-9.

[27] Khorrami M, Thawani R, Madabhushi A. Combination of peri- and intratumoral radiomic features on baseline CT scans predicts response to chemotherapy in lung adenocarcinoma. Radiat Artif Intell 2015;1.

[28] Castellano T, Moore KN, Holmlan L. An overview of immune checkpoint inhibitors in gynecologic cancers. Clin Ther 2018;40:372–8. https://doi.org/10.1016/j.clinthera.2018.01.005.

[29] Liang Y, Li W, Zhang X, Li B. Tumor-infiltrating CDS+ and FOXP3+ lymphocytes before and after neoadjuvant chemotherapy in cervical cancer. Diagn Pathol 2018;13:1–8. https://doi.org/10.1186/s13000-018-0770-4.

[30] Konishi I, Ninakubo M, Mandai M, Tsuruta Y, Kataoka N, Nagata Y, et al. Tumor response to neoadjuvant chemotherapy correlates with the expression of P-glycoprotein and PCNA but not GST-γ in the tumor cells of cervical cancer. Gynecol Oncol 1998;70:365–71. https://doi.org/10.1006/gyno.1998.5077.

[31] Ferrandina G, Ranellotto FD, Legge F, Gessi M, Salutari V, Distefano MG, et al. Prognostic role of the ratio between cytochrome-p450-2C2 in tumor and stroma compartments in cervical cancer. Clin Cancer Res 2004;10:317–23. https://doi.org/10.1158/1078-0432.CCR-03-1909-3.

[32] Gupta S, Maheshwar A, Parab P, Mahantshetty U, Hawaldar R, Sastri S, et al. Neoadjuvant Chemotherapy Followed by Radical Surgery Versus Concomitant Chemotherapy Followed by Radical Surgery Versus Concomitant Chemotherapy Plus Neoadjuvant Chemotherapy Versus Concomitant Chemotherapy Versus Concomitant Chemotherapy Versus Concomitant Chemotherapy Versus Concomitant Chemotherapy Versus Concomitant Chemotherapy Versus Concomitant Chemotherapy Plus Neoadjuvant Chemotherapy for Uterine Cervical Cancer: A Controlled Randomized Trial. J Clin Oncol 2018;36:1548–55. https://doi.org/10.1200/JCO.2017.75.9985.

[33] Vallieres M, Freeman CK, Skamene SR, El Naqia I. A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 2015;60:5471–96. https://doi.org/10.1088/0031-9155/60/19/5471.

[34] Vallieres M, Kay-Rivest E, Perrin LI, Liem X, Furstoss C, Aerts HJWL, et al. Radiomics strategies for risk assessment of tumor failure in head-and-neck cancer. Sci Rep 2017;7:10117. https://doi.org/10.1038/s41598-017-10375-1.

[35] Chen W, Chen S, Wu K, Hsieh T, Liang J, Hung Y. Prediction of local relapse and distant metastasis in patients with definitive chemoradiotherapy-treated cervical cancer by deep learning from [ 18 F ]-fluorodeoxyglucose positron emission tomography / computed tomography. Eur Radiol 2019. https://doi.org/10.1007/s00330-019-06026-x.