Original Research

Novel artificial intelligence machine learning approaches to precisely predict survival and site-specific recurrence in cervical cancer: A multi-institutional study

Chenyan Guo\textsuperscript{a,b,1}, Jue Wang\textsuperscript{a,b,1}, Yongming Wang\textsuperscript{c,1}, Xinyu Qu\textsuperscript{a,b}, Zhiwen Shi\textsuperscript{a,b}, Yan Meng\textsuperscript{a,b}, Junjun Qiu\textsuperscript{a,b,*}, Keqin Hua\textsuperscript{a,b,*}

\textsuperscript{a} Department of Gynecology, Obstetrics and Gynecology Hospital, Fudan University, 419 Fangxia Road, Shanghai 200011, China
\textsuperscript{b} Shanghai Key Laboratory of Female Reproductive Endocrine-Related Diseases, 413 Zhoushou Road, Shanghai 200011, China
\textsuperscript{c} Shanghai Changjiang Science and Technology Development Co. LTD, China

\textbf{A R T I C L E  I N F O}

\textbf{Keywords:}
Cervical cancer
Survival prediction
Machine learning
Artificial intelligence

\textbf{A B S T R A C T}

\textbf{Background:} Machine learning (ML) has been gradually integrated into oncologic research but seldom applied to predict cervical cancer (CC), and no model has been reported to predict survival and site-specific recurrence simultaneously. Thus, we aimed to develop ML models to predict survival and site-specific recurrence in CC and to guide individual surveillance.

\textbf{Methods:} We retrospectively collected data on CC patients from 2006 to 2017 in four hospitals. The survival or recurrence predictive value of the variables was analyzed using multivariate Cox, principal component, and K-means clustering analyses. The predictive performances of eight ML models were compared with logistic or Cox models. A novel web-based predictive calculator was developed based on the ML algorithms.

\textbf{Results:} This study included 5112 women for analysis (268 deaths, 343 recurrences): (1) For site-specific recurrence, larger tumor size was associated with local recurrence, while positive lymph nodes were associated with distant recurrence. (2) The ML models exhibited better prognostic predictive performance than traditional models. (3) The ML models were superior to traditional models when multiple variables were used. (4) A novel predictive web-based calculator was developed and externally validated to predict survival and site-specific recurrence.

\textbf{Conclusion:} ML models might be a better analytic approach in CC prognostic prediction than traditional models as they can predict survival and site-specific recurrence simultaneously, especially when using multiple variables. Moreover, our novel web-based calculator may provide clinicians with useful information and help them make individual postoperative follow-up plans and further treatment strategies.

\textbf{Research in context}

\textbf{Evidence before this study}

Accurate and personalized prognosis prediction of cervical cancer is required to detect early recurrence and optimize the postoperative follow-up plan. Traditionally, logistic and Cox regression models have been used as the mainstay survival analyses for oncologic research; however, they are incapable of dealing with non-linear correlations and processing big data in clinical practice. Moreover, we now lack effective tools to predict high-risk recurrence sites and guide appropriate screening. Under this situation, physicians can only assign individuals into crude categories as low- or high-risk groups without accurately accounting for the specifics of each unique patient. Therefore, a user-friendly, individual-based model that can accurately predict individual survival and site-specific recurrence simultaneously is strongly needed.

\textbf{Abbreviations:} CC, cervical cancer; FIGO, International Federation of Gynecology and Obstetrics; LVS, lymphovascular space invasion; RFS, recurrence-free survival; OS, overall survival; HPV, human papillomavirus; LEEP, loop electrosurgical excision procedure; DSI, depth of stromal invasion; LN, lymph node; NCCN, National Comprehensive Cancer Network; PCA, Principal component analysis; SVM, Support vector machine; DNN, deep neural network; DT, decision tree; RF, random forest; RSF, Random survival forest; GBDT, gradient boosting decision tree; ADASYN, adaptive synthetic sampling; C-index, concordance index; MAE, mean absolute error; AUC, area under the curve; HRs, Hazard ratios; CIs, confidence intervals.

\textsuperscript{*} Corresponding authors at: Department of Gynecology, Obstetrics and Gynecology Hospital, Fudan University, 419 Fangxia Road, Shanghai 200011, China.

\textit{E-mail addresses:} qiujunjun1113@163.com (J. Qiu), huakeqin@fudan.edu.cn (K. Hua).

\textsuperscript{1} These authors contributed equally to this article.

https://doi.org/10.1016/j.tranon.2021.101032
Received 29 November 2020; Received in revised form 24 January 2021; Accepted 28 January 2021
1936-5233/© 2021 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Added value of this study

Machine learning models have recently been considered to be useful analytic approaches for oncologic research. In this study of women with cervical cancer, we applied various machine learning methods to develop a model that can accurately predict individual risk of survival, the conditional risk of site-specific recurrence, and the specific time of recurrence-free survival or overall survival. The machine learning model exhibited better performance than traditional models, especially when using multiple variables. The results were validated using a cohort of 5112 patients from four hospitals, which is likely the largest sample size to date. To better apply the model into clinical use, we then built a web-based predictive calculator (available on https://aicer.fckyy.org.cn).

Implications of all the available evidence

Our machine learning predictive model can help doctors identify patients who are at high risk of postoperative recurrence or death, remind them of high-risk recurrence sites, and estimate recurrence-free survival or overall survival time period. Our web-based predictive calculator can provide clinicians with useful information for treatment decision-making and follow-up plan formulation.

Methods

This retrospective multicenter cohort study was approved by the Institutional Ethics Committee of Fudan University Obstetrics and Gynecology Hospital (2019-87). This study was registered in the Chinese Clinical Trial Registry (ChiCTR1900028702).

Patients

We identified 5112 patients with CC who underwent surgical resection from January 2006 to December 2017 in four tertiary hospitals as the study population. The inclusion criteria for this study included patients with pathologically confirmed stage IA1 (LVSI) to IB2 CC with complete resection. The exclusion criteria included patients with a history of prior malignancy, a preexisting history of chemotherapy or radiotherapy for other conditions, and death due to surgical complications.

Clinical information

For the eligible patients, patient demographics, laboratory test results, therapeutic data, tumor characteristics, and survival outcomes were collected from medical records. All records were reviewed simultaneously by three experts and were independently checked by two experts to ensure accuracy. The demographic variables included age and comorbidity (hypertension/diabetes). The laboratory test results included human papillomavirus (HPV) infection status. The therapeutic data included surgical approach, operative time, blood loss, transfusion, history of loop electrosurgical excision procedure (LEEP), and adjuvant treatment. The tumor characteristics included stage; tumor size; histology; depth of stromal invasion (DSI); LVSI; surgical margin; parametrial involvement; lymph node (LN) status; keratinization; differentiation; and P53, P16, and Ki67 expression.

The primary outcomes were RFS and OS. RFS was defined as the interval from the initial CC diagnosis to the first finding of any recurrence or the last follow-up. OS was defined as the interval from the initial diagnosis to CC-related death or the last follow-up. Patients who failed to reach survival events at the last follow-up were censored. Local recurrences were defined by pathologic proof of cancer in the vagina/cervix, which was confined to the pelvis, or an imaging study showing the growth of the tumor or an enlargement of any pelvic LN. Distant recurrences were defined as any recurrence outside of the pelvis including peritoneal spread or the involvement of supraclavicular LNs, the lung, the liver, the brain, etc. based on pathologic, cytologic, or radiologic evidence. The definition of local or distant recurrence was determined according to the lesions detected at the time of the first relapse after a complete workup.

According to the National Comprehensive Cancer Network (NCCN) guidelines, the preoperative workup for patients with suspicious symptoms includes history, physical examination, cervical cytologic screening, routine blood tests (including platelets), liver and renal function, electrocardiography (ECG), and imaging examinations. Radiologic imaging included chest X-ray, pelvic computed tomography (CT)/magnetic resonance imaging (MRI), or combined positron emission tomography (PET)-CT, as indicated. Cone biopsy was performed if the cervical biopsy was inadequate to define invasiveness or if an accurate assessment of microinvasive disease was required. For patients older than 60 years, echocardiography, pulmonary function tests, and urodynamic tests were also performed.

The patients were treated with adjuvant treatment after radical hysterectomy when they met one of the following two criteria: (a) patients who presented any one of several high-risk factors (surgical margin, parametrial involvement, and LN metastasis) and (b) patients who satisfied the Seldis et al. criteria for intermediate-risk factors (tumor size, LVSI, and DSI). After hospital discharge, the patients received regular follow-up in accordance with the NCCN guidelines. HPV, liquid-based cytology (LCT), tumor markers, and ultrasonography were per-
formed every 3 months for the first 2 years, every 6 months for 3 to 5 years, and annually after 5 years. Chest CT scans, upper abdominal CT scans with enhancement, and pelvic MRI were performed annually. We also performed telephone follow-up and suggested that patients who had clinical symptoms undergo imaging tests. For suspected organ or LN metastasis diagnosed by ultrasound, other imaging tests (MRI, CT, or PET/CT scan) were usually performed, and needle aspiration biopsy was conducted when necessary. The median follow-up time was 102 (36–168) months.

**Statistical analysis**

Continuous variables are reported as medians with interquartile ranges (IQRs) or means with standard deviations (SDs). Categorical variables are reported as numbers and proportions. The collinearity of all variables was evaluated using correlation matrices, and no significant interactions were identified. The associations of variables with RFS and OS were evaluated using Cox proportional hazards regression models. Variables with a P value less than 0.05 in univariate analysis were entered into multivariate survival analysis (backward selection) to identify independent predictors. The proportional hazards assumption of Cox regression was tested. Principal component analysis (PCA) and clustering analysis were also performed to further explore the relationships between clinicopathologic factors and survival outcomes. PCA was used for dimensionality reduction and feature extraction. Both the variance contribution and cumulative variance contribution were calculated to determine the number of principal components. K-means clustering analysis was performed based on the results of PCA. Hazard ratios (HRs) are presented with 95% confidence intervals (CIs).

To examine the clinicopathologic prognostic factors across traditional models (Cox and logistic models) and machine learning models, 4 datasets were examined for each separate model. Set 1 represents all basic characteristics (22 variables), including age; comorbidity; HPV infection status; surgical approach; operative time; blood loss; transfusion; history of LEEP; adjuvant treatment; FIGO stage; tumor size; histology; DSI; LVSI; surgical margin; parametrical involvement; LN status; keratinization; differentiation; and PS3, P16, and Ki67 expression; Set 2 represents 19 statistically significant variables in univariate Cox analysis (Supplementary Table 1); Set 3 represents 13 variables related to tumor characteristics, including FIGO stage; tumor size; histology; DSI; LVSI;
surgical margin; parametral involvement; LN status; keratinization; differentiation; and P53, P16, and Ki67 expression; and Set 4 represents 7 statistically significant variables in multivariate Cox analysis (Table 2).

Eight machine learning models were developed for survival prediction as a novel approach (Fig. 1). Support vector machine (SVM), deep neural network (DNN), decision tree (DT), random forest (RF), XGBoost, and LightGBM were used to predict the individual risk of recurrence/death and were compared with the traditional logistic model. Random survival forest (RSF, https://github.com/sebp/scikit-survival) and gradient boosting decision tree (GBDT) were used to predict individual specific times of RFS/OS and were compared with the traditional Cox regression model.

To train and test the generalization performance of these models, the whole dataset from four tertiary hospitals was randomly split into training and test sets (8:2) using stratified random sampling, which can ensure the consistency between patients who experienced events (recurrence/death) and those who had not. To improve the class balance in our datasets, the adaptive synthetic sampling (ADASYN) algorithm was applied to the training set [7]. To determine the optimal model parameters and avoid overfitting, we adopted the method of 5-fold cross validation (CV) based on grid search in the evaluation of the model training performance. The performance metrics of all models are the average performance metrics in the 5 validation sets.

The concordance index (C-index) and mean absolute error (MAE) were used to evaluate the performance of the models that predict individual RFS and OS. The area under the curve (AUC), sensitivity, and specificity were used to evaluate the performance of models predicting survival or recurrence probabilities. All performance parameters were calculated with 5-fold CV and an external test set.

Statistical analysis was performed in SPSS (version 21.0; SPSS Inc., Chicago, IL, USA), R 3.4.3 (Vienna, Austria; http://www.R-project.org/), and Python 3.7 (https://www.python.org/). A web-based predictive calculator was developed using Python. All tests were two-sided, and P<0.05 was considered statistically significant.

Results

Analysis of patient baseline characteristics

A total of 5842 patients with CC met the inclusion criteria. Of them, 730 patients who had incomplete medical records or were lost to follow-up were excluded. The detailed clinicopathologic characteristics of the remaining 5112 patients are listed in Table 1. The median age was 47.7 years, and 2989 (58.5%) patients underwent adjuvant treatment. Most tumors were stage I (74.5%) and of the squamous histologic type (81.7%). The median follow-up time was 102 (36–168) months. There were 343 (6.71%) women who experienced recurrence and 268 (5.24%) who died during the follow-up time. There were 179 (52.2%) patients with initial recurrence in the local region, 71 (20.7%) in the thoracic region, 43 (12.5%) in the abdominal region, 31 (9%) in the bone, and 19 (5.8%) in other regions (brain, bladder, and supraclavicular LN metastasis).

Analysis of univariate analysis, 15 variables were significantly associated with OS and recurrence, including age, FIGO stage, HPV status, adjuvant treatment, history of LEEP, tumor size, histology, DSI, LVSI, surgical margin, parametral involvement, LN status, keratinization, differentiation, and immunohistochemistry (P53, P16, and Ki67) (Supplementary Table 1). Factors significant in univariate analysis regarding the specific recurrence site are shown in Supplementary Table 2. It is worth noting that FIGO stage, parametral involvement, LN status, and adjuvant treatment were associated with recurrence in any specific site based on the univariate analysis.

Table 1 Baseline characteristics of stage IA1(LVSI)- IB2 cervical cancer patients.

| Characteristics                  | Number of patients (n = 5112) |
|----------------------------------|-------------------------------|
| **Clinical variables**           |                               |
| Age, years                       | 477 (± 9.6)                   |
| FIGO stage                       |                               |
| IA1 (LVSI)                       | 27 (0.5)                      |
| IA2                              | 136 (2.7%)                    |
| IB1                              | 3202 (62.6%)                  |
| IB2                              | 605 (11.8%)                   |
| IA1                              | 734 (14.4%)                   |
| IA2                              | 338 (6.6%)                    |
| IB1                              | 39 (0.8%)                     |
| IB2                              | 31 (0.6%)                     |
| Comorbidity                      |                               |
| Yes                              | 768 (15%)                     |
| No                               | 4344 (85%)                    |
| HPV infection                    |                               |
| Yes                              | 1963 (38.4%)                  |
| No                               | 594 (11.6%)                   |
| Unknown                          | 2555 (50%)                    |
| Adjuvant treatment               |                               |
| Yes                              | 2989 (58.5%)                  |
| No                               | 2123 (41.5%)                  |
| **Surgery related variables**    |                               |
| Surgery approach                 |                               |
| MH                               | 4040 (79%)                    |
| LMH                              | 3799 (74.3%)                  |
| RMH                              | 236 (4.6%)                    |
| Trans-vaginal                    | 5 (0.1%)                      |
| OH                               | 1072 (21%)                    |
| Operative time, min              | 213.5 (165, 251)              |
| Blood loss, ml                   | 335.9 (150,400)               |
| Transfusion                      |                               |
| Yes                              | 369 (7.2%)                    |
| No                               | 4743 (92.8%)                  |
| LEEP                             | 982 (19.2%)                   |
| No                               | 4130 (80.8%)                  |
| **Pathologic variables**         |                               |
| Tumor size, cm                   |                               |
| >0.5                             | 975 (19.1%)                   |
| [0.5,1)                          | 96 (1.9%)                     |
| [1.1,5)                          | 338 (6.6%)                    |
| [1.5,2)                          | 428 (8.4%)                    |
| [2,2.5)                          | 422 (8.3%)                    |
| [2,5,3)                          | 518 (10.1%)                   |
| [3,3.5)                          | 727 (14.2%)                   |
| [3,5,4)                          | 623 (12.2%)                   |
| [4,4.5)                          | 375 (7.3%)                    |
| [4,5,5)                          | 188 (3.7%)                    |
| ≥5                               | 422 (8.3%)                    |
| Histology                        |                               |
| SCC                              | 4179 (81.7%)                  |
| AC                               | 576 (11.3%)                   |
| AS                               | 281 (5.5%)                    |
| Rare type                        | 76 (1.5%)                     |
| DSI                              |                               |
| Negative                         | 1219 (23.8%)                  |
| Inner 2/3                        | 1542 (30.2%)                  |
| Outer 1/3                        | 2131 (42.6%)                  |
| LVSI                             |                               |
| Yes                              | 2129 (41.6%)                  |
| No                               | 2983 (58.4%)                  |
| Surgical margin                  |                               |
| Yes                              | 399 (7.8%)                    |
| No                               | 4713 (22.2%)                  |
| Parametral involvement           |                               |
| Yes                              | 255 (5%)                      |
| No                               | 4857 (95%)                    |
| LN metastasis                    |                               |
| Yes                              | 1000 (19.6%)                  |
| Pelvic LNs                       | 710 (13.9%)                   |
| Common iliac LNs                 | 248 (4.9%)                    |
| Para-aortic LNs                  | 42 (0.8%)                     |
| No                               | 4112 (80.4%)                  |

(continued on next page)
Table 1 (continued)

| Characteristics | Number of patients (n = 5112) |
|-----------------|--------------------------------|
| Keratinization  |                                |
| Yes             | 1168 (22.8%)                   |
| No              | 2012 (39.4%)                   |
| Non-SCC         | 533 (18.3%)                    |
| Unknown         | 999 (19.5%)                    |
| Differentiation |                                |
| Low             | 105 (2.1%)                     |
| Intermediate    | 232 (4.5%)                     |
| High            | 31 (0.6%)                      |
| Unknown         | 4744 (92.8%)                   |
| PS3             |                                |
| Negative        | 1642 (32.1%)                   |
| +               | 2215 (43.3%)                   |
| ++              | 76 (1.5%)                      |
| +++             | 26 (0.5%)                      |
| ++++            | 2 (0%)                         |
| Unknown         | 1151 (22.5%)                   |
| P16             |                                |
| Negative        | 213 (4.2%)                     |
| +               | 2909 (56.9%)                   |
| +++             | 583 (11.4%)                    |
| ++++            | 58 (1.1%)                      |
| Unknown         | 1010 (19.8%)                   |
| Ki67            |                                |
| Negative        | 15 (0.3%)                      |
| 0–20%           | 517 (10.1%)                    |
| 20–40%          | 997 (19.5%)                    |
| 40–60%          | 1038 (20.3%)                   |
| 60–80%          | 1147 (22.4%)                   |
| 80–100%         | 389 (7.6%)                     |
| Unknown         | 1009 (19.7%)                   |
| Follow-up, months | 90 (18–162)                   |

Multivariate Cox analysis showed that (1) the following seven variables were independent predictors of both OS and recurrence: FIGO stage, adjuvant therapy, tumor size, histology, DSI, parametrical involvement, and LN status (Table 2). (2) The following seven variables were independent predictors of local recurrence: FIGO stage, adjuvant therapy, tumor size, histology, DSI, surgical margin, and parametrical involvement. (3) The following five variables were independent predictors of thoracic recurrence: FIGO stage, adjuvant therapy, histology, parametrical involvement, and LN status. (4) The following three variables were independent predictors of abdominal recurrence: FIGO stage, parametrical involvement, and LN status. (5) The following two variables were independent predictors of bone recurrence: FIGO stage and LN status (Table 3). Notably, for site-specific recurrence, larger tumor size was associated with local recurrence, while positive LNs were associated with distant recurrence based on the multivariate Cox analysis.

To further investigate the associations between multiple variables and patient prognosis, PCA and clustering analysis were performed. After applying one-hot encoding on all variables, the first 40 principal components explained more than 85% of the total variance (Supplementary Table 3). Clustering analysis was then performed based on the results of PCA, and two prognosis-related clinical phenotypes (groups A and B, which represent good prognosis and poor prognosis, respectively) were determined according to the elbow method (Fig. 2). Group B had significantly worse RFS (HR 3.863, 95% CI 2.508–5.95) and OS (HR 5.987, 95% CI 3.317–10.808) than group A (Fig. 3). Compared to patients in group A, patients in group B had significantly higher FIGO stages, more comorbidities, a higher frequency of LEEP, and positive HPV. In addition, larger tumor sizes, non-squamous cell carcinoma, deeper stromal invasion, LVI, positive surgical margins, positive parametrical involvement, and LN metastasis were more common in group B (Supplementary Table 4).

Collectively, based on univariate analysis, multivariate analysis, PCA, and clustering analysis, we identified certain potential variables associated with patient prognosis, which were selected for the following prognostic model development.

Comparison of the prognostic predictive performance between machine learning models and traditional models in 4 datasets

To predict the individual risk of recurrence or death, 6 machine learning algorithms were tested in 4 datasets (22, 19, 13, and 7 variables in set 1, set 2, set 3, and set 4, respectively) and were compared with the logistic model. The average AUC value, sensitivity, and specificity obtained from all machine learning models and the logistic model in both the validation group and test group are presented in Table 4 and 5. For predicting the risk of recurrence, SVM exhibited the best performance with higher accuracy, sensitivity, and specificity values than the logistic model in 4 datasets. For estimating the individual risk of death, the best predictive performance was obtained by RF. Regarding the specific recurrence site (local or distant recurrence), we again found that the AUC of SVM significantly surpassed those of the remaining models.

To predict the specific time of RFS and OS, we selected RSF and GBDT as the machine learning methods for their superiority in time prediction. The comparison was then made between these two models and the Cox regression model in 4 datasets (Table 6). The C-index of the two machine learning models was markedly higher than that of the Cox model in all 4 datasets. Similar findings were observed for MAE. These findings suggest that machine learning models showed better predictive performance with a higher C-index and lower MAE than the Cox regression model.

To evaluate model performance with different variables, we compared machine learning models and the Cox model across 4 datasets (22, 19, 13, and 7 variables in set 1, set 2, set 3, and set 4, respectively) for predicting RFS and OS (Table 6). The Cox model exhibited the best performance in dataset 4 (7 significant variables in multivariate analysis) but the worst performance in dataset 1 (all 22 variables). In contrast, the performance of machine learning models improved as more variables were added to the models. These findings indicate that the prognostic predictive performance of machine learning models outperformed that of traditional logistic or Cox models, especially when using multiple variables.

Establishment of a novel web-based predictive calculator based on the machine learning models

To better apply the prediction models in clinical practice and create user-friendly access, the statistical formulas were implemented in a web-based predictive calculator. After entering the clinicopathologic information of the patient and time of current follow-up after surgery, physicians/users can estimate the patient’s individual conditional risk of death, risk of recurrence, risk of site-specific recurrence, RFS, and OS. For example, as the screenshot shows in Fig. 4, the estimated conditional probabilities of overall death, overall recurrence, local recurrence, and distant recurrence were 10.2%, 11.17%, 2.68%, and 2.16%, respectively. This calculator may help physicians identify patients who are at high risk of recurrence or death, remind them of
Table 2
Multivariate Cox analysis factors associated with recurrence-free survival and overall survival in stage IA1 (LVI) to IIIB cervical cancer patients.

| Characteristics          | No. | Multivariate |                |                |                |
|--------------------------|-----|--------------|----------------|----------------|----------------|
|                          |     | RFS          | OS             |                |                |
|                          |     | HR (95%CI)   | P              | HR (95%CI)     | P              |
| FIGO (%)                 |     | <0.001       | 1              | <0.001         | 1              |
| IA1 (LVI)                | 27  | 1            | 1.005 [0.136,7.347] | 1.057 [0.137,5.682] | 0.041         |
| IA2                      | 136 | 0.502 [0.081,1.105] | 0.735 [0.176,3.073] | 0.862 [0.233,7.246] | 0.009         |
| IB1                      | 3202| 0.633 [0.196,2.044] | 1.678 [0.47,0.537] | 1.299 [0.233,7.246] | 0.001         |
| IB2                      | 605 | 1.019 [0.312,3.326] | 1.347 [0.319,5.682] | 1.213 [0.425,10.71] | 0.001         |
| IA1                      | 734 | 1.137 [0.349,3.708] | 1.828 [0.431,7.759] | 1.299 [0.233,7.246] | 0.001         |
| IA2                      | 338 | 1.506 [0.458,4.949] | 1.299 [0.233,7.246] | 1.213 [0.425,10.71] | 0.001         |
| IB1                      | 39  | 1.062 [0.249,4.530] | 1.299 [0.233,7.246] | 1.213 [0.425,10.71] | 0.001         |
| IB2                      | 31  | 1.444 [0.368,5.667] | 1.299 [0.233,7.246] | 1.213 [0.425,10.71] | 0.001         |

Adjuvant therapy (%)<0.001 0.041
No 2123 1 1
Yes 2989 1.51 [1.109,2.055] 1.44 [1.016,2.041] 0.001

Tumor size, cm<0.001 0.001
<0.5 975 1 1
[0.5,1) 96 0.517 [0.188,1.423] 0.59 [0.188,1.856] 0.001
[1,1.5) 338 0.458 [0.217,0.968] 0.373 [0.157,0.899] 0.001
[1.5,2) 428 1.873 [1.053,3.332] 1.701 [0.867,3.337] 0.001
[2,2.5) 422 0.886 [0.448,1.753] 0.552 [0.233,1.309] 0.001
[2.5,3) 518 0.956 [0.526,1.736] 0.862 [0.432,1.719] 0.001
[3,3.5) 727 0.884 [0.481,1.625] 0.838 [0.416,1.687] 0.001
[3.5,4) 623 1.141 [0.637,2.042] 1.054 [0.537,2.071] 0.001
[4,4.5) 375 0.9 [0.439,1.844] 0.723 [0.311,1.685] 0.001
[4.5,5) 188 1.33 [0.684,2.588] 1.257 [0.589,2.682] 0.001
≥5 422 1.808 [0.948,3.448] 1.488 [0.696,3.138] 0.001

Histology<0.001
SCC 4179 1 1
AC 576 1.706 [1.241,2.345] 1.921 [1.351,2.733] 0.001
AS 281 1.69 [1.142,2.506] 1.949 [1.272,2.991] 0.001
Rare type 76 2.395 [1.341,4.277] 2.134 [1.062,4.29] 0.001
DSI 0.006
Negative 1219 1 1
<2/3 1542 1.007 [0.627,1.618] 1.435 [0.771,2.668] 0.001
≥2/3 2351 1.609 [1.032,2.507] 2.414 [1.336,4.36] 0.001

Parametrical involvement<0.001
No 4857 1 1
Yes 255 2.034 [1.509,2.742] 1.851 [1.324,5.29] 0.001

LN metastasis<0.001
No 4112 1 1
Pelvic LNs 710 1.414 [1.066,1.876] 1.513 [1.096,2.809] 0.001
Common iliac LNs 248 3.078 [2.263,4.186] 3.5 [2.494,4.911] 0.001
Para-aortic 42 4.503 [2.523,8.04] 6.543 [3.485,12.285] 0.001

Fig. 3. Survival outcome comparisons between group A and B. Recurrence-free survival (A); Overall survival (B).

Discussion

In this study, 5112 women from four tertiary hospitals were included in the analysis. There were 268 deaths and 343 recurrences during the follow-up period of 102 (36–168) months, of which 179 were local recurrences and 164 were distant recurrences. (1) Based on multivariate analysis for site-specific recurrence, we found that larger tumor sizes were associated with local recurrence, while positive LNs were associated with distant recurrence. (2) The machine learning models exhibited a better performance than traditional logistic or Cox regression models in estimating prognostic outcomes. With regard to the prediction of the individual risk of recurrence and death, the best results were obtained

high-risk recurrence sites and make individual postoperative follow-up plans and further treatment strategies for CC patients (available on https://aicer.fckyy.org.cn).

C. Guo, J. Wang, Y. Wang et al. Translational Oncology 14 (2021) 101032
by SVM. For predicting the specific time of RFS and OS, all machine learning models outperformed the Cox regression model. (3) The machine learning models were superior to traditional models when using multiple variables, and the performance of the machine learning models improved as more variables were added to the models. (4) A novel simple and efficient user-friendly web-based calculator was developed and externally validated to precisely predict postoperative survival and site-specific recurrence in CC patients.

Tumor recurrence after surgical resection remains a challenge in treating CC patients. A major barrier to the effective prevention of postoperative recurrence is the inability to identify “at risk” and “high-risk” recurrence sites. Traditionally, the estimation of risk was based on clinicians’ experience and knowledge by assigning individuals into crude categories as low- or high-risk groups without accurately accounting for the specifics of each unique patient. Considering these shortcomings, the development of accurate clinical models to predict an individual’s future risk of site-specific recurrence and death is urgently needed and will be an effective prevention approach and will also greatly optimize follow-up strategies.

In a review of the previous literature, most studies applied logistic or Cox proportional hazards regression models with nomograms to predict oncologic prognostic outcomes [3,4]. However, because of the inability of these models to address nonlinear relationships, which occur in real-world settings, an increasing number of studies have started to apply artificial intelligence and machine learning in the prediction of survival outcomes [8-10]. As a novel analytic approach, machine learning models are able to automatically learn feature characteristics from raw data, fit censored survival data, and exhibit better performance when processing larger datasets and dealing with nonlinear relationships between variables. However, only a few studies have integrated machine learning methods into the prediction of oncologic survival outcomes in the area of CC. The sample sizes of these studies were relatively small;
## Table 4.
Comparison of model performance (probability prediction of recurrence and survival, happen or not).

| Model                | Recurrence          | Survival            |          |
|----------------------|---------------------|---------------------|----------|
|                      | Validation group (n = 1023) | Test group (n = 1023) |          |
|                      | AUC | Sen | Spe  | AUC | Sen | Spe  |          |
| Logistic             | 0.701(0.016) | 0.727(0.039) | 0.675(0.014) | 0.785 | 0.725 | 0.679 |          |
| SVM                  | 0.703(0.016) | 0.769(0.033) | 0.636(0.011) | 0.794 | 0.768 | 0.659 |          |
| ANN                  | 0.853(0.022) | 0.739(0.037) | 0.768(0.009) | 0.728 | 0.561 | 0.749 |          |
| DT                   | 0.685(0.017) | 0.857(0.015) | 0.515(0.109) | 0.607 | 0.678 | 0.445 |          |
| RF                   | 0.845(0.041) | 0.876(0.072) | 0.814(0.015) | 0.741 | 0.522 | 0.874 |          |
| XGBoost              | 0.778(0.025) | 0.881(0.056) | 0.740(0.020) | 0.751 | 0.667 | 0.674 |          |
| LightGBM             | 0.897(0.051) | 0.879(0.110) | 0.915(0.011) | 0.757 | 0.464 | 0.929 |          |

**Set 2**

| Model                | Validation group (n = 1023) | Test group (n = 1023) |          |
|----------------------|-----------------------------|------------------------|----------|
| Logistic             | 0.699(0.026) | 0.719(0.053) | 0.679(0.014) | 0.783 | 0.696 | 0.684 |          |
| SVM                  | 0.701(0.021) | 0.762(0.042) | 0.939(0.015) | 0.790 | 0.783 | 0.655 |          |
| ANN                  | 0.697(0.024) | 0.758(0.068) | 0.637(0.057) | 0.652 | 0.536 | 0.710 |          |
| DT                   | 0.712(0.029) | 0.728(0.059) | 0.696(0.055) | 0.645 | 0.580 | 0.681 |          |
| RF                   | 0.823(0.031) | 0.883(0.060) | 0.763(0.010) | 0.752 | 0.623 | 0.780 |          |
| XGBoost              | 0.803(0.030) | 0.844(0.055) | 0.762(0.010) | 0.744 | 0.594 | 0.791 |          |
| LightGBM             | 0.811(0.031) | 0.858(0.059) | 0.764(0.010) | 0.747 | 0.623 | 0.747 |          |

**Set 3**

| Model                | Validation group (n = 1023) | Test group (n = 1023) |          |
|----------------------|-----------------------------|------------------------|----------|
| Logistic             | 0.688(0.024) | 0.677(0.059) | 0.688(0.020) | 0.802 | 0.783 | 0.714 |          |
| SVM                  | 0.696(0.020) | 0.711(0.043) | 0.680(0.019) | 0.803 | 0.797 | 0.708 |          |
| ANN                  | 0.718(0.030) | 0.750(0.053) | 0.667(0.068) | 0.728 | 0.681 | 0.644 |          |
| DT                   | 0.649(0.009) | 0.869(0.012) | 0.430(0.025) | 0.699 | 0.855 | 0.431 |          |
| RF                   | 0.740(0.033) | 0.738(0.051) | 0.743(0.020) | 0.751 | 0.609 | 0.767 |          |
| XGBoost              | 0.679(0.015) | 0.801(0.046) | 0.556(0.022) | 0.764 | 0.754 | 0.594 |          |
| LightGBM             | 0.804(0.038) | 0.833(0.062) | 0.776(0.018) | 0.766 | 0.609 | 0.766 |          |

**Set 4**

| Model                | Validation group (n = 1023) | Test group (n = 1023) |          |
|----------------------|-----------------------------|------------------------|----------|
| Logistic             | 0.634(0.030) | 0.679(0.054) | 0.590(0.008) | 0.752 | 0.725 | 0.605 |          |
| SVM                  | 0.682(0.018) | 0.760(0.029) | 0.604(0.020) | 0.778 | 0.768 | 0.612 |          |
| ANN                  | 0.644(0.024) | 0.780(0.790) | 0.508(0.049) | 0.764 | 0.754 | 0.630 |          |
| DT                   | 0.659(0.034) | 0.663(0.117) | 0.655(0.059) | 0.732 | 0.725 | 0.624 |          |
| RF                   | 0.855(0.035) | 0.872(0.063) | 0.837(0.112) | 0.738 | 0.754 | 0.835 |          |
| XGBoost              | 0.664(0.032) | 0.739(0.057) | 0.591(0.012) | 0.770 | 0.696 | 0.798 |          |
| LightGBM             | 0.647(0.026) | 0.734(0.065) | 0.560(0.047) | 0.780 | 0.754 | 0.566 |          |
to our knowledge, the largest dataset (n = 768) was reported by Matsuo et al. [11]. Additionally, the majority of these studies were conducted in single institutions and lacked external validation [11–13]. Their main prediction outcomes were the probability of recurrence or death, but they seldom took specific recurrence sites (local or distant recurrence) into consideration, which is essential for planning appropriate follow-up strategies.

Considering the deficiencies of previous studies and the promising application value of machine learning algorithms, in the present study, we applied 6 machine learning methods to predict the risk of recurrence/death and 2 machine learning methods to predict the specific time of RFS(OS) based on the data of 5112 CC patients, and we externally validated these models using data from four tertiary hospitals. In our study, we observed that the machine learning models exhibited superior prognostic predictive performance compared to linear regression models (logistic and Cox models) in estimating individual risk of recurrence/death or RFS/OS. It is worth noting that the machine learning models also outperformed the traditional logistic model in predicting site-specific recurrence (local or distant metastasis). In addition, the performance of the machine learning models improved when more features were added to the models. These findings revealed that machine learning might be a better analytic approach in prognostic prediction, especially when using multiple variables.

Admittedly, although many studies have identified clinicopathologic predictors for OS and recurrence after radical hysterectomy of CC, additional information on site-specific recurrence is still needed to select diagnostic procedures during surveillance. In the current study, we identified FIGO stage, adjuvant therapy, tumor size, histology, DSI, parametrical involvement, and LN status as independent predictors of OS and recurrence using multivariate analysis. For site-specific recurrence, larger tumor size was an independent predictor of local recurrence in multivariate analysis but not a significant predictor of distant recurrence. This finding was consistent with those of previous studies which reported that tumor size was strongly correlated with local recurrence rather than distant recurrence [14–17]. In addition, consistent with the findings of certain published studies [16,18,19], we found that positive LNs were independent predictors of distant recurrence (thoracic, abdominal, and bone recurrence), whereas they were not significant for local recurrence. Collectively, our study supported the identification of risk factors for recurrence in each specific site, which also ensures the rationality for building models to estimate site-specific recurrence.

Of note, in this study, we were the first group to establish a web-based calculator that inputs the clinicopathologic features of individual patients into the developed machine learning algorithms. Clinicians can estimate the conditional risk of death and site-specific recurrence and generate personalized surveillance strategies accordingly, including when to follow up patients and what strategies to use for further diagnosis and treatment. For example, if a patient shows a high recurrence risk in the local region, then the clinician can recommend pelvic MRI for effective screening. If a patient is estimated to have a high thoracic recurrence risk, then chest CT might be a better choice. This can greatly save medical resources and optimize individualized surveillance in precision medicine.
There are several limitations in this study. First, considering the retrospective nature of this study, future prospective studies are still warranted. Second, our models were developed and validated based on Chinese patients, and the generalizability needs further validation with non-Chinese patient data. Third, as the follow-up time was relatively short (<5 years), caution should be taken in applying this model to estimate long-term prognostic outcomes.

In conclusion, this study used machine learning technology as a novel approach to develop prediction models to precisely estimate survival and site-specific recurrence in CC patients. We trained and externally validated the models based on the data of 5112 CC patients from four tertiary hospitals, which is the largest multicenter cohort to date. Our models can provide multitask prediction using various machine learning methods, which can estimate the individual probability of overall survival, recurrence, and site-specific recurrence as well as RFS and OS times at the same time. The machine learning models outperformed traditional logistic or Cox models, especially when using multiple variables. Of note, we built a novel user-friendly web-based calculator based on our machine learning algorithms for the first time. This calculator can help to identify patients who are at high risk of postoperative recurrence or death and provide clinicians with useful information for treatment decision-making and follow-up plan formulation.

Contributors

Conceptualization: Chenyan Guo, Junjun Qiu,
Methodology: Chenyan Guo, Yue Wang, Yongming Wang
Data acquisition: Chenyan Guo, Xin Yu Qu, Zhiwen Shi, Yan Meng
Validation: Keqin Hua, Junjun Qiu
Writing-original draft: Chenyan Guo
Writing-review & editing: all authors
Supervision: Keqin Hua, Junjun Qiu

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This study received aid from the National Natural Science Foundation of China (No. 81971361; to Jun-jun Qiu), the Natural Science Foundation of Shanghai Science and Technology (No. 19ZR1406900; to Jun-jun Qiu), the Research and Innovation Project of Shanghai Municipal Education Commission (No. 2019-01-07-00-07-E00050; to Ke-qin Hua), the artificial Intelligence Innovation Project of Shanghai Municipal Commission of Economy and Informatization (No. 2018-RGZN-02041; to Ke-qin Hua), the Shanghai research project of Shanghai Municipal Commission of Economy and Informatization (No. 2018-RGZN-02041; to Ke-qin Hua), the Shanghai Shenkang Hospital Development Center Funding (No. SHDC2020CR4087; to Jun-jun Qiu); the Shanghai Municipal Health Commission (No. 202040498; to Jun-jun Qiu); the Shanghai Shenkang Hospital Development Center Funding (No. SHDC2020CR1045B; to Keqin Hua); and the clinical research project of Shanghai Municipal Health Commission (No.202040498; and Shanghai “Rising Stars of Medical Talent” Youth Development Program (No. AB83030002019004; to Jun-jun Qiu).

Data sharing statement

The data that support the findings of this study are available upon request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 68 (6) (2018) 394–424.
[2] W. Chen, R. Zheng, P.D. Baade, S. Zhang, H. Zeng, F. Bray, et al., Cancer statistics in China, 2015, CA Cancer J. Clin. 66 (2) (2016) 115–152.
[3] S. Polteauer, C. Grimm, H. Hofstetter, N. Concin, C. Natter, A. Sturzda, et al., Nomogram prediction for overall survival of patients diagnosed with cervical cancer, Br. J. Cancer 107 (6) (2012) 918–924.
[4] H. Zhou, X. Li, Y. Zhang, Y. Jia, T. Hu, R. Yang, et al., Establishing a nomogram for stage IIA-IIIB cervical cancer patients after complete resection, Asian Pac. J. Cancer Prevent.: APJCP 16 (9) (2015) 3773–3777.
[5] W.J. Koh, B.E. Greer, N.R. Abs-Rastam, S.M. Apte, S.M. Campos, K.R. Cho, et al., Cervical cancer, version 2.2015, J. Natl. Compr. Cancer Netw.: JNCCN 13 (4) (2015) 395–404 quiz.
[6] A. Seidlis, B.N. Bundy, M.Z. Rotman, S.S. Lenz, L.I. Maderspach, R.J. Zaino, A randomized trial of pelvic radiation therapy versus no further therapy in selected patients with stage IB carcinoma of the cervix after radical hysterectomy and pelvic lymphadenectomy: a gynecologic oncology group study, Gynecol. Oncol. 73. (2) (1999) 177–183.
[7] H. Ye, Y. Bai, E.A. Garcia, S. Li, ADASYN: adaptive synthetic sampling approach for imbalanced learning’, in: Proceedings of the International Joint Conference on Neural Networks, 2008, pp. 1322–1328.
[8] A.V. Karhade, Q. Thio, P. Ogink, J. Kim, S. Lozano-Calderon, K. Raskin, et al., Development of machine learning algorithms for prediction of 5-year spinal chordoma survival, World Neurosurg. 119 (2018) e842–e847.
[9] M. Montazeri, M. Montazeri, A. Beigzadeh, Machine learning models in breast cancer survival prediction. Technology and health care, Off. J. Eur. Soc. Eng. Med. 24 (1) (2016) 31–42.
[10] P. Gupta, S.F. Chiang, P.K. Sahoo, S.K. Mohapatra, J.F. You, D.D. Othoni, et al., Prediction of colon cancer stages and survival period with machine learning approach, Cancers 11 (12) (2019).
[11] K. Matzu, S. Purushotham, B. Jiang, R.S. Mandalbaum, T. Takuchi, Y. Liu, et al., Survival outcome prediction in cervical cancer: Cox models vs deep-learning model, Am. J. Obstet. Gynecol. 220 (4) (2019) 38.1.e1–e14.
[12] B. Obrutz, M. Kusy, A. Szemczuk, M. Obrutz, J. Kluska, Prediction of 5-year overall survival in cervical cancer patients treated with radical hysterectomy using computational intelligence methods, BMC Cancer 17 (1) (2017) 840.
[13] B. Obrutz, M. Kusy, A. Szemczuk, M. Obrutz, J. Kluska, Prediction of 10-year overall survival in patients with operable cervical cancer using a probabilistic neural network, J. Cancer 10 (18) (2019) 4189–4195.
[14] S. Beriwal, N. Kannan, P. Sukumvanich, S.D. Richard, J.L. Kelley, R.P. Edwards, et al., Complete metabolic response after definitive radiation therapy for cervical cancer: patterns and factors predicting for recurrence, Gynecol. Oncol. 127 (2) (2012) 303–306.
[15] I.L. Atahan, C. Onal, E. Ozayr, F. Yiliz, U. Selek, F. Kose, Long-term outcome and prognostic factors in patients with cervical carcinoma: a retrospective study, Int. J. Gynecol. Cancer: Off. J. Int. Gynecol. Cancer Soc. 17 (4) (2007) 833–842.
[16] A.C.M. Queiroz, V. Fabri, H. Mantoan, S.M. Sanches, A.P.G. Guimarães, A.R.G. Ribeiro, et al., Risk factors for pelvic and distant recurrence in locally advanced cervical cancer, Eur. J. Obstet. Gynecol. Reprod. Biol. 235 (2019) 6–12.
[17] G. Comerici, B.S. Bolger, G. Flannelly, M. Maini, A. de Barros Lopes, J.M. Monaghan, Prognostic factors in surgically treated stage IB-IIIB carcinoma of the cervix with negative lymph nodes, Int. J. Gynecol. Cancer: Off. J. Int. Gynecol. Cancer Soc. 8 (1) (1998) 23–26.
[18] R. Ma, H. Zhang, L. Zou, Y. Qu, Outcome and prognostic factors of distant metastasis in patients with advanced cervical squamous cell carcinoma treated with concurrent chemoradiotherapy, Zhonghua Fu Chan Ke Za Zhi 50 (2) (2015) 125–130.
[19] L. Xin, W. Li-Chun, Z. Ying, Z. Li-Na, L. Wei-Wei, P. Li-Jian, D. Yun-Zhi, H. Jing, S. Mei, The prognosis and risk stratification based on pelvic lymph node characteristics in patients with locally advanced cervical squamous cell carcinoma treated with concurrent chemoradiotherapy, Int. J. Gynecol. Cancer: Off. J. Int. Gynecol. Cancer Soc. 27 (2) (2017) 409.