CLINICAL ARTICLE

Gynecology

Endometrial Cancer Individualized Scoring System (ECISS): A machine learning-based prediction model of endometrial cancer prognosis

Sherif A. Shazly[1] | Pluvio J. Coronado[2] | Ercan Yılmaz[3] | Rauf Melekoglu[3] | Hanifi Sahin[3] | Luca Giannella[4] | Andrea Ciavattini[4] | Giovanni Delli Carpini[4] | Jacopo Di Giuseppe[4] | Angel Yordanov[5] | Konstantina Karakadieva[5] | Nevena Milenova Nedelcheva[5] | Mariela Vasileva-Slaveva[5] | Juan Luis Alcazar[6] | Enrique Chacon[6] | Nabil Manzour[6] | Julio Vara[6] | Erbil Karaman[7] | Onur Karaaslan[7] | Latif Hacıoğlu[7] | Duygu Korkmaz[7] | Cem Onal[8] | Jure Knez[9] | Federico Ferrari[10] | Esraa M. Hosni[11] | Mohamed E. Mahmoud[11] | Gena M. Elassall[11] | Mohamed S. Abdo[11] | Yasmin I. Mohamed[11] | Amr S. Abdelbadie[12] | Middle-Eastern College of Obstetricians and Gynaecologists (MCOG) Muti-Center Studies (MCS) office and Artificial Intelligence Unit (AI)

[1]Women's Service, Leeds Teaching Hospitals NHS Trust, Leeds, UK

[2]Department of Obstetrics and Gynecology, Hospital Clinico San Carlos, Madrid, Spain

[3]Department of Obstetrics and Gynecology, Inonu University, Malatya, Turkey

[4]Woman's Health Sciences Department, Gynecologic Section, Polytechnic University of Marche, Ancona, Italy

[5]Department of Gynecologic Oncology, Medical University of Pleven, Pleven, Bulgaria

[6]Gynecologic Oncologic Division, Clinica Universidad de Navarra, Navarra, Spain

[7]Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Van Yuzuncu Yıl University, Van, Turkey

[8]Department of Radiation Oncology, Baskent University Faculty of Medicine, Adana, Turkey

[9]University Medical Center Maribor, Maribor, Slovenia

[10]Department of Clinical and Experimental Sciences, University of Brescia, Brescia, Italy

Abstract

Objective: To establish a prognostic model for endometrial cancer (EC) that individualizes a risk and management plan per patient and disease characteristics.

Methods: A multicenter retrospective study conducted in nine European gynecologic cancer centers. Women with confirmed EC between January 2008 to December 2015 were included. Demographics, disease characteristics, management, and follow-up information were collected. Cancer-specific survival (CSS) and disease-free survival (DFS) at 3 and 5 years comprise the primary outcomes of the study. Machine learning algorithms were applied to patient and disease characteristics. Model I: pretreatment model. Calculated probability was added to management variables (model II: treatment model), and the second calculated probability was added to perioperative and postoperative variables (model III).

Results: Of 1150 women, 1144 were eligible for 3-year survival analysis and 860 for 5-year survival analysis. Model I, II, and III accuracies of prediction of 5-year CSS were 84.88%/85.47% (in train and test sets), 85.47%/84.88%, and 87.35%/86.05%, respectively. Model I predicted 3-year CSS at an accuracy of 91.34%/87.02%. Accuracies of models I, II, and III in predicting 5-year DFS were 74.63%/76.72%, 77.03%/76.72%, and 80.61%/77.78%, respectively.

[11]Middle-Eastern College of Obstetricians and Gynaecologists, Leeds, UK

[12]Department of Obstetrics and Gynecology, Aswan University, Aswan, Egypt

Correspondence
Sherif A. Shazly, Women's Service, Leeds Teaching Hospitals NHS Trust, Leeds, LS9 7TF, UK.
Email: sherif.shazly@nhs.net

Conclusion: The Endometrial Cancer Individualized Scoring System (ECISS) is a novel machine learning tool assessing patient-specific survival probability with high accuracy.

KEYWORDS
Artificial intelligence, Disease-free survival, Overall survival, Uterine cancer

1 | INTRODUCTION

Endometrial cancer (EC) is the most common gynecologic cancer in low- and middle-income countries and is the sixth most diagnosed malignancy in women worldwide. In 2020, 417 367 new cases were diagnosed and 97 370 EC-related mortalities were reported globally.[1] In most instances, EC is diagnosed early when tumor growth is confined to the uterine cavity. The 5-year overall survival of EC, in the absence of metastasis, ranges between 74% and 91%.[2]

Ideally, a cancer staging system would be designed to classify disease spread, indicate prognosis, and contribute to treatment decision.[3] With regard to EC, FIGO (the International Federation of Gynecology & Obstetrics) staging has been adopted as the standard classification system since 1970.[4] FIGO staging is generally reflective of disease prognosis; a Stage III EC is associated with 5-year overall survival of 58%, compared with 91% with Stage I disease.[3] However, it is not inclusive of all prognostic factors, such as histologic type, pathologic grade, and lymphovascular space invasion.[5] In fact, disease grade, and myometrial and cervical stromal invasion were associated with greater prognostic impact than pelvic lymph node metastasis.[6] These factors would impact prognosis and therapeutic decision making, and contribute to management plan over staging alone.[7] Multifactorial interactions add to the complexity of EC classification and highlight the need for an individualized scoring system for the assessment of disease prognosis and delineation of treatment.

Machine learning is a subtype of artificial intelligence where the machine progressively recognizes patterns that link variables to a specific outcome. In comparison to traditional statistics, machine learning establishes comprehensive individualized prediction models rather than recognizing associations between individual variables and a clinical outcome.[8] The objective of this study is to develop an EC scoring system that predicts individualized prognosis based on patient demographics and disease characteristics using machine learning algorithms. In addition, the study aims to predict prognosis in response to a therapeutic plan, which would facilitate individualized treatment decisions in women with EC.

2 | MATERIALS AND METHODS

The Endometrial Cancer International Database (ECID) is a multi-centric data collection project that was launched by the Middle-East Obstetrics and Gynaecology Graduate Education foundation for the purpose of the current study. ECID comprises the collaboration of nine European cancer centers. Women who were diagnosed with EC and received treatment in the contributing centers between January 1, 2008 and December 31, 2015 were considered eligible. Inclusion criteria encompassed women aged 18 years or older, with confirmed diagnosis of EC, and who were followed up for at least 3–5 years unless mortality was reported earlier. Women were excluded from the study if there was inadequate documentation of management/pathology of the disease, if they were lost to follow up before 3 years post-treatment, and if they were diagnosed with synchronous cancers. Women included in the study should have provided an authorization to use their deidentified data for research purposes. The study received ethics committee approval under number aswu/ 530/5/21 on May 10, 2020 from the hosting institute (Aswan University Faculty of Medicine Ethical Committee). Each contributing center received ethical approval from the respective ethics committees and institutional review board approval was given by all centers contributing to the study before data collection. Participants did not have to provide informed consent as data was retrospectively collected and available to use for research purposes.

Using relevant codes, EC patients, treated within the intended period, were identified. A standardized EXCEL spreadsheet was designed for the study and shared among participating centers. Variables of interest included patient demographics (e.g. patient age, parity, body mass index, ethnicity, major medical comorbidities), preoperative assessment of disease extent per clinical assessment and imaging (e.g. tumor size, extent of myometrial involvement, parametrial invasion, cervical and vaginal invasion, pelvic and para-aortic lymph node metastasis, omental/peritoneal involvement, lung and liver metastasis, pleural effusion, ascites), disease characteristics (FIGO staging, histopathologic type, tumor grade, lymphovascular space invasion [LVSI]), and genetic predisposition (positive family history, Lynch syndrome). Management details were also collected; surgery information incorporated type of hysterectomy, whether oophorectomy and omentectomy/omental biopsy were performed, and lymph node management if any (sentinel lymph nodes, pelvic lymphadenectomy, para-aortic lymphadenectomy, or lymph node sampling). Data on adjuvant treatment were also considered including type of radiotherapy, total dose and number of fractions, and duration of chemotherapy. Intraoperative and postoperative complications, residual disease after surgery, duration of follow up, incidence and site of recurrence, overall mortality, and disease-specific mortality were identified. Collected data did not include any

identifiable information. Primary outcome of the study was cancer-specific survival (CSS) rate at 3 and 5 years after initiation of treatment. Disease-free survival (DFS) rate at 3 and 5 years constituted the secondary outcome.

2.1 | Data pre-processing

Data from contributing centers were merged and reviewed for integrity and consistency. As part of data pre-processing, all variables were converted to either continuous or categorical values that corresponded to categorical values defined in the study protocol. Variables were described using mean, median, standard deviation, range, and percentage depending on variable type and distribution.

Missing data were assessed, and imputation of missed values was considered to preserve data size. Two approaches were initially tested: MICE (Multiple Imputation by Chained Equations) and KNN (k-Nearest Neighbors). Both approaches use machine learning algorithms to predict missing values and are superior to traditional methods of management of missing values.[9] To determine the selected strategy, both approaches were tested on "virtual missing values". "Virtual missing values" were created by deleting 30% of the values of three categorical and continuous variables that yielded the most missing data. MICE and KNN were applied separately to impute virtual missing values and imputed values were compared to actual values to test accuracy. The three selected variables were Lynch syndrome, number, and size of enlarged para-aortic lymph nodes on preoperative imaging. Actual missing values were 39.8%, 40%, and 40%, respectively. Following imputation, accuracy of prediction of missing values of Lynch syndrome was 97.95% with both approaches. For the number of enlarged para-aortic lymph nodes, mean standard error was used to assess accuracy and it was 0.1 with both methods. However, mean standard errors for para-aortic lymph node size were 0.9 and 4.8 with MICE and KNN, respectively. Accordingly, MICE was deemed superior and was used to treat missing values in the ECID database.

Feature extraction using principal component analysis was applied for dimensionality reduction before running the algorithm. This was applied to model I, as discussed later, due to the large number of variables.

2.2 | Machine learning approach

Data were defined as Xi and Yi where Xi represents independent variables (features) and Yi represents dependent variables (target). Four algorithms were run (logistic regression, Support vector machine, Xgboost, and random forest) for each target to select the highest-performing model. Performances of the four models were overall comparable. However, for purposes of simplicity, we will refer to the Xgboost model in the Results section because it performed relatively better than other models. Train: test split was applied to data using a 0.8:0.2 ratio, where the train set was used

to develop the model and the test set was used to evaluate its performance.

Three models were created to predict each of the four outcomes (CSS and DFS at 3 years and 5 years). Model I (preoperative model) included patient characteristics and preoperatively determined disease characteristics. This model aims to determine the individualized probability of survival in percentage at the time of disease diagnosis as an alternative/adjuvant to FIGO staging (score I). Hence, FIGO staging was not included in the model. Calculated probability of survival from model I (score I) was included as a feature in model II, in addition to details of management (surgery, radiotherapy, chemotherapy). The aim of this model (management model) was to determine prognosis per selected management plan(s). Model III (postoperative model) used probability from model II (score II) in addition to features determined after surgery (e.g. residual disease and treatment complications) to revise probability of survival following surgery. The term "score" in all models refers to probability of survival as a percentage. Features included in each model are summarized in Table S1.

2.3 | Model evaluation

Model performance was appraised using model accuracy, F1 score, and precision and recall scores. The F1 score equals $2 \times ([precision \times recall]/[precision + recall])$ and is the sum of the predictive performance of the model. Precision is equivalent to positive predictive value (Precision = true positives/[true positives + false positives]), whereas recall represents true positive / (true positive + false negative). Learning curves were plotted using train and validation sets, created by k-fold cross validation, to rule out over-fitting and ensure plateauing of the learning process over cohort size (Figure S1). Feature importance was used to graph the contribution of different features in predicting outcome and was used for both model II and model III. Feature importance was not used in model I because principal component analysis was used in this model. Calculated probabilities from each model were plotted against each outcome using the receiver operating characteristics curve to test the diagnostic performance of these models against the prognostic performance of FIGO staging. A box and whisker plot was used to visually compare the distribution of the three scores among survivors and non-survivors. For conventional statistics, a P value less than 0.05 was used as a cut-off for significance.

Machine learning models were created using PYTHON 3.8, through the ANACONDA 3 interface. Conventional statistical analysis was performed using STATA software, version 14 (STATA Corp., College Station, TX, USA).

3 | RESULTS

Out of 1150 women who were initially included, 1091 were eligible for 3-year survival analysis and 860 for 5-year survival analysis. The difference between the two values represented women who

were lost to follow up after 3 years after management was initiated (199 patients) and those who were censored for non-cancer-related deaths (32 patients) (Figure S2). Mean age of study population was 63.87 ± 10.99 years, and 680 (62.3%) women were white. The most common medical comorbidities in the study population were hypertension (477, 43.8%) and diabetes mellitus type 2 (266, 24.4%). Diagnosis of Lynch syndrome was confirmed in 23 (2%) women. Five hundred and sixty-two (51.51%) women were diagnosed with EC at Stage Ia and 260 (23.83%) were diagnosed at Stage Ib. Endometroid EC was present in 942 (86.3%) women, followed by papillary serous EC (61, 5.6%). EC was classified as grade 1 in 513 (47%) women and grade 3 in 270 (24.8%) women. In 279 (24.4%) women, LVSI was identified. Patient and disease characteristics are summarized in Table S2. Eight hundred and thirty-nine (73.3%) women underwent class I hysterectomy and 177 (15.5%) had class II hysterectomy. Brachytherapy was received by 434 (38%) women and 388 (33.9%) were treated with external beam radiotherapy. Treatment and follow-up details are shown in Table S3.

3.1 | 5-year cancer-specific survival

Model I (preoperative model) predicted 5-year CSS with an accuracy of 84.9% in the train set and 85.5% in the test set. The F1 score, precision score, and recall score of this model were 91, 86.3, and 96.2, respectively. Accuracy of model II (management model) was 85.5% and

84.9% in the train and test sets, respectively. F1 score of this model was 91.4, precision score was 86.2, and recall score was 97.6. The most contributing features to this model were score I, total dose of external beam radiotherapy, and adjuvant chemotherapy (Figure 1). Calculating probability of survival using model II, the predicted 5-year CSS (score II) was found to change per management plan. For example, score II in patients with Stage Ia, grade 2 endometrioid EC and positive LVSI was 0.85 ± 0.25 in those who did not receive brachytherapy versus 0.96 ± 0.09 in women who were treated with brachytherapy. Table S4 shows examples of score II in different clinical scenarios per optional management options. Model III (postoperative model) was associated with an accuracy of 87.4% (train set) and 86.1% (test set) and yielded F1 score, precision score, and recall score of 92.3, 86.6, and 99.5, respectively (Table 1). Score II and pelvic lymph node metastasis contributed the most to this model (Figure 1).

Diagnostic performances of scores I, II, and III were plotted against 5-year CSS. Areas under the curve (AUC) were 0.87 (95% confidence interval [CI] 0.85–0.90), 0.90 (95% CI 0.88–0.92), and 0.93 (95% CI 0.91–0.95) for the three scores, respectively. The AUC of FIGO staging was 0.75 (95% CI 0.71–0.79) (Figure 2). The distribution of the three scores among survivors and non-survivors is illustrated in Figure 3.

Score I, II, and III were strongly correlated. Correlation coefficient (r) between scores I and II, scores I and III, and scores II and III were 0.92, 0.82, and 0.90, respectively. However, 11.3% of women had greater than 15% difference between score I and II (Figure S3).
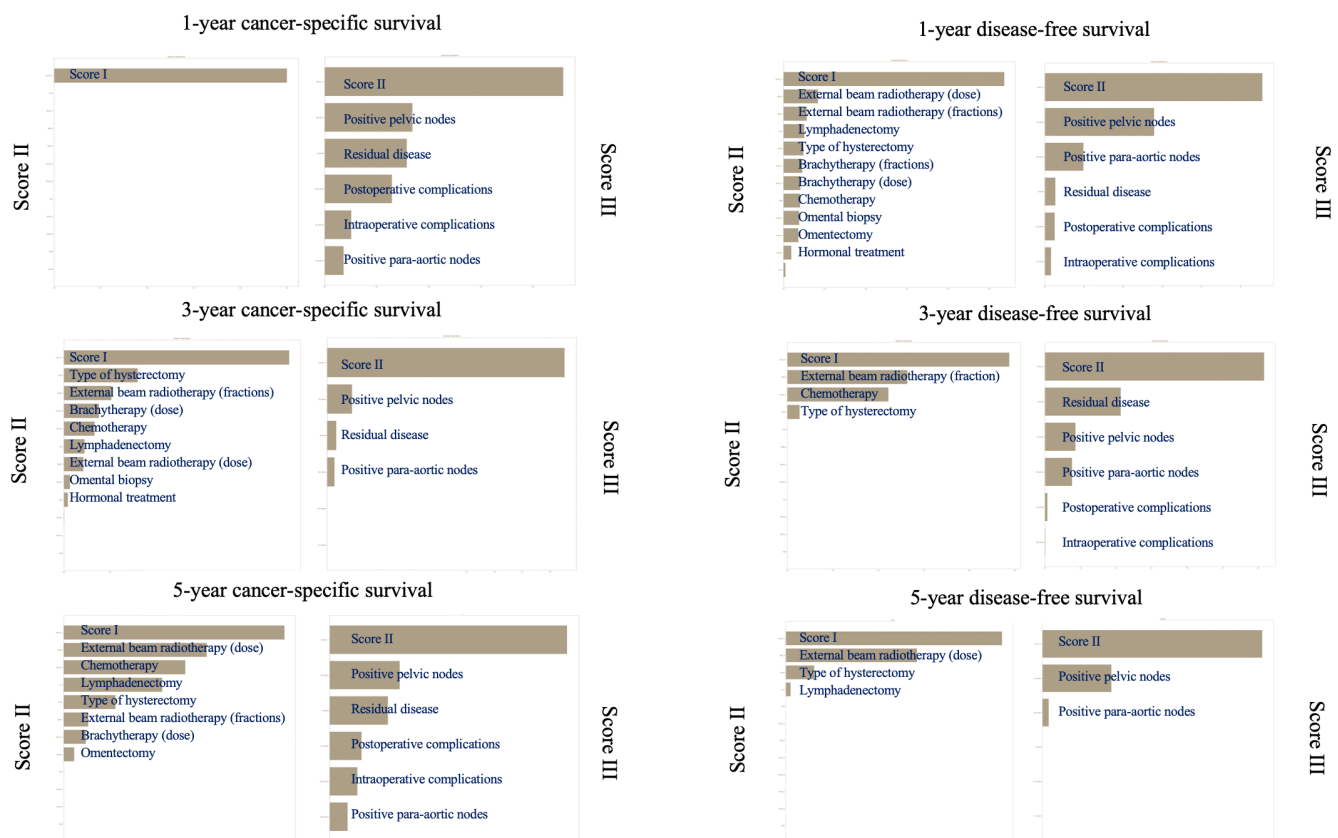

FIGURE 1 Feature importance in predicting cancer-specific survival and disease-free survival.

TABLE 1 Performance of machine learning prediction models of cancer-specific survival and disease-free survival rates

	Logistic regression				Support vector machine				Random forest				XGBoost			
	Model acc. Train/Test	F1 score	P score	R score	Model acc. Train/Test	F1 score	P score	R score	Model acc. Train/Test	F1 score	P score	R score	Model acc. Train/Test	F1 score	P score	R score
Cancer-specific survival																
At 1 y																
Score I	98.17/95.43	99.07	98.38	99.77	97.94/96.35	98.96	97.93	100.0	97.82/97.26	98.90	97.82	100.0	97.82/97.26	98.90	97.82	100.0
Score II	97.94/96.35	98.96	97.93	100.0	97.82/96.80	98.90	97.82	100.0	97.82/97.26	98.90	97.82	100.0	97.82/97.26	98.90	97.82	100.0
Score III	97.82/97.26	98.89	97.82	100.0	97.82/97.26	98.89	97.82	100.0	97.82/97.26	98.89	97.82	100.0	97.82/97.26	98.89	97.82	100.0
At 3 y																
Score I	91.23/87.50	95.23	91.82	98.92	90.87/86.54	95.07	91.17	99.32	88.70/85.58	94.01	88.70	100.0	91.34/87.02	95.33	91.42	99.59
Score II	91.45/86.54	95.62	92.31	99.19	90.38/86.54	94.84	90.42	99.73	88.70/85.58	94.01	88.70	100.0	92.90/87.02	96.14	92.81	99.73
Score III	88.70/85.58	94.01	88.70	100.0	88.94/85.58	94.13	88.92	99.86	88.70/85.58	94.01	88.70	100.0	91.70/86.06	95.52	91.66	99.73
At 5 y																
Score I	84.30/83.14	90.49	86.97	94.31	84.59/84.30	90.73	86.64	95.23	80.96/81.40	89.27	80.62	100.0	84.88/85.47	90.97	86.33	96.15
Score II	85.03/83.72	91.00	86.83	95.59	85.17/84.88	91.04	87.35	95.05	82.41/83.72	89.79	83.13	97.24	85.47/84.88	91.38	86.17	97.61
Score III	86.77/84.88	92.05	87.83	96.70	83.92/83.72	92.08	88.49	95.96	79.22/80.81	88.40	79.22	100.0	87.35/86.05	92.57	86.58	99.45
Disease-free survival																
At 1 y																
Score I	94.00/92.34	96.91	94.00	100.0	94.34/92.34	97.08	94.32	100.0	94.00/92.34	96.91	94.00	100.0	95.36/93.24	97.59	95.30	100.0
Score II	94.12/92.79	96.97	94.11	100.0	95.36/92.79	97.59	95.30	100.0	94.00/92.34	96.91	94.00	100.0	95.59/93.24	97.71	95.52	100.0
Score III	94.00/92.34	96.91	94.00	100.0	94.46/92.34	97.14	94.43	100.0	94.00/92.34	96.91	94.00	100.0	94.00/92.34	96.91	94.00	100.0
At 3 y																
Score I	84.62/84.19	90.96	86.12	96.37	84.03/84.19	90.71	85.11	97.10	80.30/81.86	90.46	80.30	100.0	83.22/84.19	89.08	83.19	99.13
Score II	84.27/85.58	90.80	85.60	96.66	83.68/83.72	91.15	87.45	93.03	80.30/81.86	89.08	80.30	100.0	84.27/84.65	90.73	86.07	95.94
Score III	85.78/84.65	91.47	88.26	94.92	86.01/85.12	91.60	91.60	94.92	80.30/81.86	89.08	80.30	100.0	86.25/85.12	91.71	88.84	94.77
At 5 y																
Score I	76.10/74.60	83.42	78.56	88.93	75.83/76.19	83.58	78.83	88.93	68.39/72.49	81.23	68.39	100.0	74.63/76.72	84.15	73.48	98.45
Score II	78.75/77.78	85.24	81.20	89.71	72.51/76.19	82.88	72.19	97.28	77.29/75.66	84.44	79.45	90.10	77.03/76.72	85.23	76.07	96.89
Score III	80.74/78.31	86.66	82.34	91.46	80.74/78.31	86.66	82.34	91.46	68.39/72.49	81.23	68.39	100.0	80.61/77.78	86.28	83.61	89.13

Abbreviations: Model acc., model accuracy; P score, precision; R score, recall score.
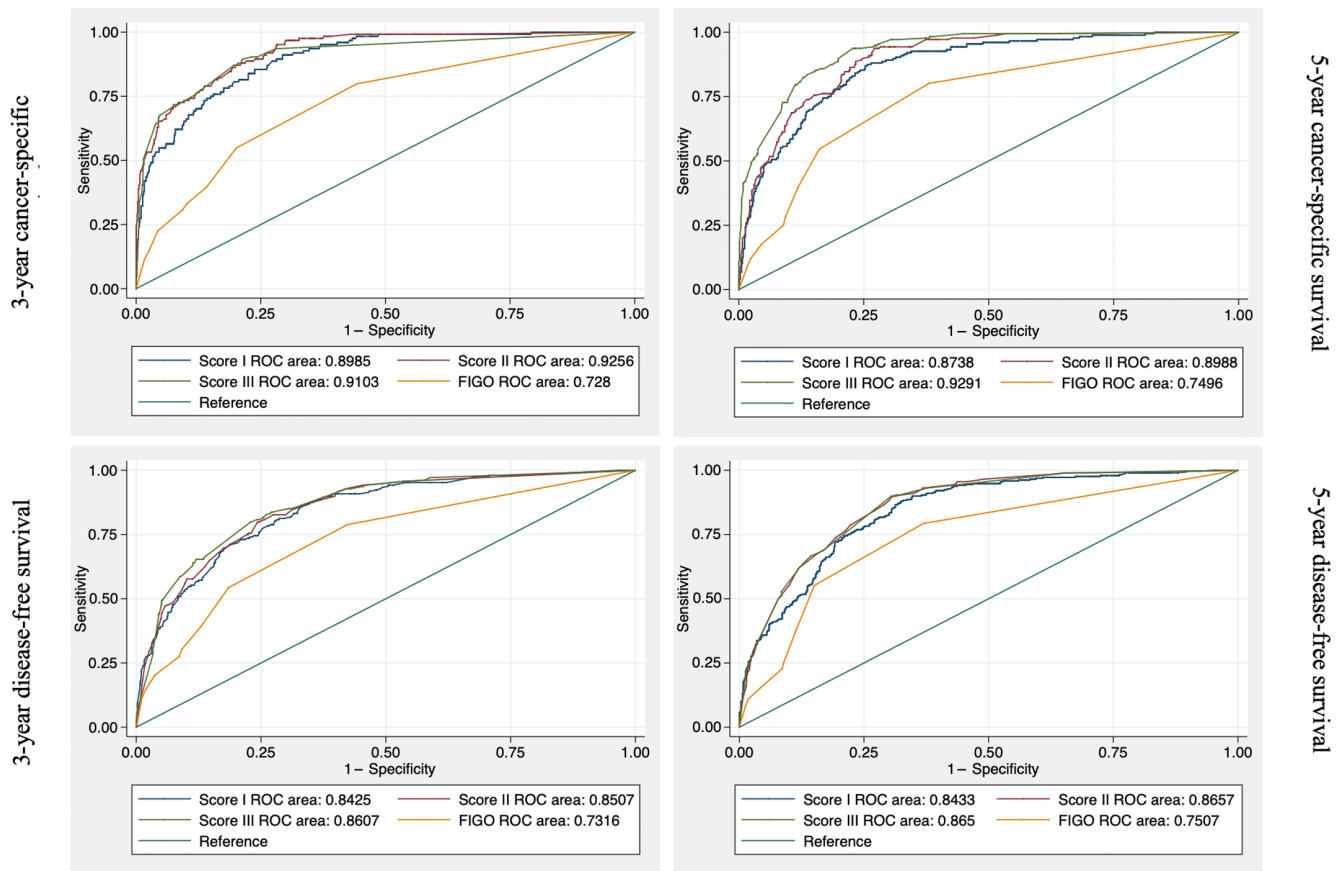
FIGURE 2   Receiver operating characteristic curves comparing prognostic performance of score I, II, and III and FIGO staging.

3.2 | 3-year cancer-specific survival

Model I, II and III prediction accuracies of 3-year CSS in the train and test sets were 91.3%/87%, 92.9%/87%, and 91.7%/86.1%, respectively. Model I was associated with F1 score of 95.3, precision score of 91.4, and recall score of 99.6. The F1 score of model II was 96.1, precision score was 92.8, and recall score was 99.7. The most significant features that contributed to model prediction were score I, type of hysterectomy, and external beam radiotherapy fractions (Figure 1). Model III had values of 95.5, 91.7, and 99.7 for F1 score, precision score, and recall score, respectively (Table 1). AUC of scores I, II and III were 0.90 (95% CI 0.87–0.92), 0.93 (95% CI 0.90–0.95), and 0.91 (95% CI 0.88–0.93) for 3-year CSS, respectively. FIGO staging yielded an AUC of 0.73 (95% CI 0.68–0.77) (Figure 2). Scores I and II were greater than 15% different in 6.5% of women.

3.3 | 5-year disease-free survival

Accuracy of model I in predicting 3-year DFS was 74.6% (train set) and 76.7% (test set). Model II accuracies were 77% and 76.7% in the train and test sets, respectively. Score I, external beam radiotherapy dose, and type of hysterectomy were the most important features

in this model (Figure 1). Model III performed at accuracies of 80.6% (train set) and 77.8% (test set). The model was most influenced by score II and lymph node metastasis (Figure 1). AUC was 0.87 (95% CI 0.82–0.87) for score I, 0.87 (95% CI 0.84–0.89) for score II, 0.84 (95% CI 0.84–0.89) for score III, and 0.75 (95% CI 0.72–0.78) for FIGO staging (Figure 2). In 7.1% of women, score I and II were greater than 15% different.

3.4 | 3-year disease-free survival

For the 3-year DFS, predictive accuracies of model I were 83.2% and 84.2% in the train and test sets, respectively. Accuracy of model II as shown in the train test was 84.3% and it was 84.7% in the test set. In terms of feature importance, score I, external beam radiotherapy fractions, and chemotherapy were the most significant, in order (Figure 1). Model III performed at accuracy of 86.3% (train set) and 85.1% (test set) (Table 1). Score II, residual disease, and lymph node metastasis were the most important features (Figure 1). Compared with FIGO staging, which showed an AUC of 0.73 (95% CI 0.69–0.77), the AUC of scores I, II, and III were 0.84 (95% CI 0.81–0.87), 0.85 (95% CI 0.82–0.88), and 0.86 (95% CI 0.83–0.89), respectively (Figure 2). The difference between scores I and II was greater than 15% in 5.1% of women.
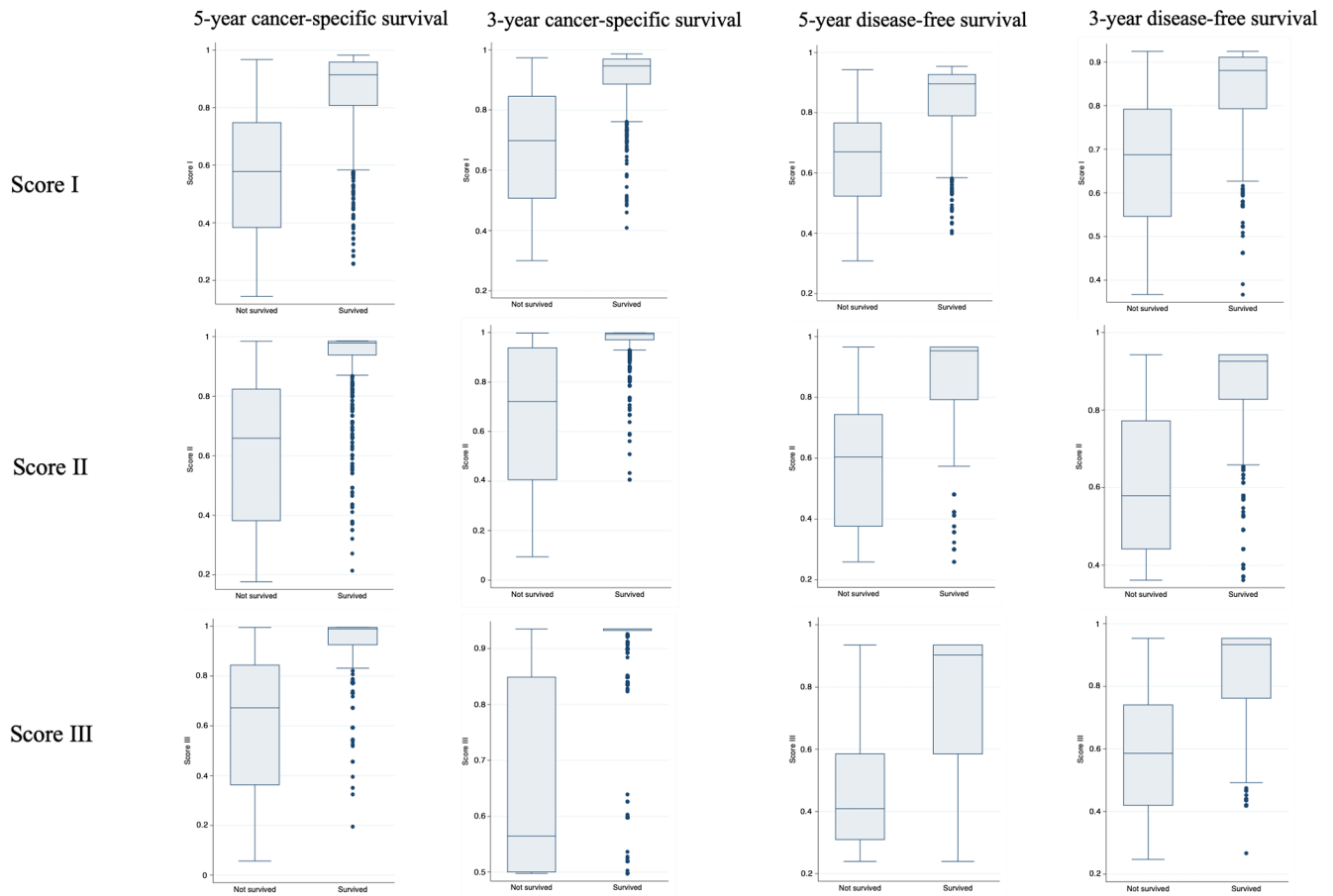
FIGURE 3 Score distribution per survival outcome.

4 | DISCUSSION

The current study established a series of machine learning models to predict DFS and CSS in women newly diagnosed with EC. These models were designed to determine individualized disease prognosis at baseline considering a wide range of patient and disease characteristics and were associated with higher diagnostic performance compared with traditional FIGO staging. In addition, these models compute individualized probability of survival in response to different management strategies before treatment. All models exhibited high accuracy and precision and were associated with convergent and plateaued learning curves.

Since the 1970s, FIGO staging has been globally recognized as the standard classification system of EC.[4] Similar to other cancer staging systems, it incorporates disease extent of invasion, metastasis to regional lymph nodes, and evidence of distant metastasis, and is used mainly to provide information on disease prognosis and to determine treatment plan.[10] Although disease stage contributes significantly to disease prognosis, emerging evidence has supported the crucial rule of a range of patient and disease characteristics in cancer outcome. Well-established prognostic factors such as histopathologic type and tumor grade are not incorporated in the current staging.[11] Similarly, LVSI was found to adversely impact overall

survival in EC patients.[12,13] Extent of tumor invasion into surrounding structures is well recognized by disease staging. Nevertheless, specific tumor size may act as an independent factor to disease survival. Canlorbe et al.[14] studied 633 women with early-stage EC and concluded that tumor size of 35 mm or more was associated with lower DFS, probably as a sequence of nodal involvement.

Prognostic impact of patient characteristics has also been investigated. Of these characteristics, age was recognized as a risk factor of poor prognosis. Son et al.[15] analyzed data of 551 women with EC who underwent primary surgery or fertility-sparing treatment. Younger age group, less than or equal to 40 years, was linked to good prognosis. A secondary analysis of a randomized clinical trial, including 173 patients with EC, showed that age over 65 years was significantly associated with poor survival (5-year CSS 92.1% in women at or younger than 65 years versus 78.4% in women over 65 years, $P < 0.001$). Interestingly, high body mass index aggravated poor prognosis only in this particular age group.[16] Hence, the present study not only highlights the impact of individual factors on disease prognosis, but also indicates possible interactions among these factors. Although these prognostic factors are not included in FIGO staging system, they are growing evidence that, at least some of these factors would influence treatment decisions. The PORTEC-2 trial highlighted the effect of brachytherapy on DFS in women with

additional risk factors, including myometrial invasion, grade 3, and age greater than 60 years.[17] LVSI was recognized by Bosse et al.[18] as the most prominent poor prognostic factor based on a pooled analysis of data from the PORTEC-1 and -2 trials and accordingly, adjuvant external beam radiotherapy and/or chemotherapy were recommended in women with stage I EC if substantial LVSI was evident. The significance of these factors has added to the complexity of the treatment decision and has warranted additional classifications within the same disease stage, proposed by national and international guidelines, to determine prognosis and proper management. The national comprehensive cancer network integrates risk factors and tumor grade within Stage I to outline management.[19] On the other hand, the European Society for Medical Oncology classifies EC into low, intermediate, high-intermediate, or high risk based on risk factors including stage, grade, extent of invasion, LVSI, and genetic factors.[20]

Consequently, there is a wide spectrum of risk factors that would contribute to the prognosis and treatment decisions, which have been variably included in contemporary guidelines. The current study proposed a series of prediction models that endorse a comprehensive panel of general and prognostic factors to individualize prognosis. The models were created using machine learning algorithms, which carry the privilege of identifying patterns and interactions among potential predictors to create robust models. Machine learning considers all variables and precludes risk of bias secondary to variable selection in traditional statistics and sufficiently treats potential collinearity.[8] The study used a multicenter large database to enhance machine learning algorithms and support the generalizability of results. Unlike FIGO staging, which is a surgical staging that should be fully determined after the procedure, the Endometrial Cancer Individualized Scoring System (ECISS) model I employs preoperative data, and model II provides the option of testing a treatment plan based on preoperative assessment.[21]

However, the present study is prone to inherent limitations due to its retrospective nature. The current study has not included immunohistochemical and genetic features, which have been increasingly considered in EC studies and have been recently highlighted by internationally recognized guidelines.[22-24] Given the fact that genetic and immunohistochemical assessments have not been fully established and standardized in all gynecologic cancer centers, the current study, which presents the first version of ECISS, did not include these variables in the models to permit generalizability. Nevertheless, the ECISS development project is a planned project that will be launched by the end of 2022 to secure an ongoing process to develop further versions of the model, which would consider more centers, larger data, and more variables including genetic and immunohistochemistry results as well as validation of the ECISS model on external data (https://www.mogge-obgyn.com/eciss-project). Unlike traditional statistics, complexity of machine learning algorithms interferes with its applicability because scores cannot be calculated using simple mathematics. Therefore, a software was created using the created

models and is available in the current link: https://www.mogge-obgyn.com/eciss-project.

In conclusion, ECISS is a novel machine learning scoring system that endorses individualized prognosis of women with EC based on patient and disease characteristics. The system is reflective of the treatment decision and can be incorporated in management planning. ECISS development is supported by an ongoing project that aims at expanding the database and input variables in future versions.

AUTHOR CONTRIBUTIONS

SAS was the study's principal investigator, contributed to study design and data analyses and was the main writer. PJC, RM, LG, AY, JLA, EK, CO, JK, and FF were center principal investigators, and contributed to study conduct, data collection and reviewing, and manuscript writing and review. EY, HS, AC, GDC, JDG, KK, NMN, MV, EC, NM, JVG, OK, LH, and DK contributed to data collection, reviewing, and cleaning, and to manuscript review. EMH coordinated the study and contributed to study registration, protocol writing, study design, and manuscript review. MEM, GME, MSA, and YIM contributed to data reviewing and pre-processing, data analysis, and manuscript writing. ASA contributed to study design, protocol writing, and to manuscript writing and review.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Sherif A. Shazly 🄳 https://orcid.org/0000-0003-1062-9227

REFERENCES

1. Endometrial cancer statistics [World Cancer Research Fund International]. https://www.wcrf.org/cancer-trends/endometrial-cancer-statistics/. Last accessed on May 24th, 2022.
2. Morice P, Leary A, Creutzberg C, Abu-Rustum N, Darai E. Endometrial cancer. *Lancet*. 2016;387(10023):1094-1108.
3. Zaino RJ. FIGO staging of endometrial adenocarcinoma: a critical review and proposal. *Int J Gynecol Pathol*. 2009;28(1):1-9.
4. Kottmeier H. Classification and staging of malignant tumors in the female pelvis. *Int J Gynecol Obstet*. 1971;9:172.
5. Uharček P. Prognostic factors in endometrial carcinoma. *J Obstet Gynaecol Res*. 2008;34(5):776-783.
6. Köbel M, Atenafu EG, Rambau PF, et al. Progesterone receptor expression is associated with longer overall survival within high-grade histotypes of endometrial carcinoma: a Canadian high risk endometrial cancer consortium (CHREC) study. *Gynecol Oncol*. 2016;141(3):559-563.
7. de Boer SM, Powell ME, Mileshkin L, et al. Adjuvant chemoradiotherapy versus radiotherapy alone in women with high-risk endometrial cancer (PORTEC-3): patterns of recurrence and post-hoc survival analysis of a randomised phase 3 trial. *Lancet Oncol*. 2019;20(9):1273-1285.
8. Shazly SA, Trabuco EC, Ngufor CG, Famuyide AO. Introduction to machine learning in obstetrics and gynecology. *Obstet Gynecol*. 2022;139(4):669-679.

9. Deng Y, Chang C, Ido MS, Long Q. Multiple imputation for general missing data patterns in the presence of high-dimensional data. *Sci Rep*. 2016;6(1):1-10.

10. Pecorelli, S., 2009. Revised FIGO staging for carcinoma of the vulva, cervix, and endometrium, Nordic Society of Gynecologic Oncology. Retrieved from https://policycommons.net/artifacts/1771180/revised-figo-staging-for-carcinoma-of-the-vulva-cervix-and-endometrium/2502826/ on 05 Jun 2022. CID: 20.500.12592/d5k17v.

11. Amant F, Moerman P, Neven P, Timmerman D, Van Limbergen E, Vergote I. Endometrial cancer. *Lancet*. 2005;366(9484):491-505.

12. Cusano E, Myers V, Samant R, et al. Prognostic significance of lymphovascular space invasion in the absence of lymph node metastases in early-stage endometrial cancer. *Int J Gynecol Cancer*. 2018;28(5):890-894.

13. Raffone A, Travaglino A, Raimondo D, et al. Lymphovascular space invasion in endometrial carcinoma: a prognostic factor independent from molecular signature. *Gynecol Oncol*. 2022;165(1):192-197.

14. Canlorbe G, Bendifallah S, Laas E, et al. Tumor size, an additional prognostic factor to include in low-risk endometrial cancer: results of a French multicenter study. *Ann Surg Oncol*. 2016;23(1):171-177.

15. Son J, Carr C, Yao M, et al. Endometrial cancer in young women: prognostic factors and treatment outcomes in women aged≤40 years. *Int J Gynecol Cancer*. 2020;30(5):631-639.

16. Panici PB, Basile S, Salerno MG, et al. Secondary analyses from a randomized clinical trial: age as the key prognostic factor in endometrial carcinoma. *Am J Obstet Gynecol*. 2014;210(4):363.e1-363.e10.

17. Nout R, Smit V, Putter H, et al. Vaginal brachytherapy versus pelvic external beam radiotherapy for patients with endometrial cancer of high-intermediate risk (PORTEC-2): an open-label, non-inferiority, randomised trial. *Lancet*. 2010;9717:816-823.

18. Bosse T, Peters EE, Creutzberg CL, et al. Substantial lymph-vascular space invasion (LVSI) is a significant risk factor for recurrence in endometrial cancer–a pooled analysis of PORTEC 1 and 2 trials. *Eur J Cancer*. 2015;51(13):1742-1750.

19. Koh W-J, Abu-Rustum NR, Bean S, et al. Uterine neoplasms, version 1.2018, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2018;16(2):170-199.

20. Santoro A, Angelico G, Travaglino A, et al. New pathological and clinical insights in endometrial cancer in view of the updated ESGO/ESTRO/ESP guidelines. *Cancer*. 2021;13(11):2623.

21. Rungruang B, Olawaiye AB. Comprehensive surgical staging for endometrial cancer. *Rev Obstet Gynecol*. 2012;5(1):28-34.

22. Markova I, Duskova M, Lubusky M, et al. Selected immunohistochemical prognostic factors in endometrial cancer. *Int J Gynecol Cancer*. 2010;20(4):576-582.

23. Concin N, Matias-Guiu X, Vergote I, et al. ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. *Int J Gynecol Cancer*. 2021;31(1):12-39.

24. Cavaliere AF, Perelli F, Zaami S, et al. Towards personalized medicine: non-coding RNAs and endometrial cancer. *Healthcare: 2021*. MDPI; 2021:965.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.