Explainable machine learning for predicting recurrence-free survival in endometrial carcinosarcoma patients

Samantha Bove[1†], Francesca Arezzo[2,3†], Gennaro Cormio[2,4], Erica Silvestris[2], Alessia Cafforio[4], Maria Colomba Comes[1]*, Annarita Fanizzi[1]*, Giuseppe Accogli[1], Gerardo Cazzato[5], Giorgio De Nunzio[6,7], Brigida Maiorano[8], Emanuele Naglieri[2], Andrea Lupo[1], Elsa Vitale[9], Vera Loizzi[2,10†] and Raffaella Massafra[1†]

[1]Laboratorio di Biostatistica e Bioinformatica, Fisica Sanitaria, I.R.C.C.S. Istituto Tumori "Giovanni Paolo II", Bari, Italy, [2]Ginecologia Oncologica, I.R.C.C.S. Istituto Tumori "Giovanni Paolo II", Bari, Italy, [3]Dipartimento di Medicina di Precisione e Rigenerativa e Area Jonica (DiMePRe-J), Università degli Studi di Bari "Aldo Moro", Bari, Italy, [4]Dipartimento Interdisciplinare di Medicina (DIM), Università degli Studi di Bari "Aldo Moro", Bari, Italy, [5]Dipartimento dell'Emergenza e dei Trapianti di Organi, Università degli Studi di Bari "Aldo Moro", Bari, Italy, [6]Laboratory of Biomedical Physics and Environment, Department of Mathematics and Physics "E. De Giorgi", Università del Salento, Lecce, Italy, [7]Advanced Data Analysis in Medicine (ADAM), Laboratory of Interdisciplinary Research Applied to Medicine (DReAM), Università del Salento, Lecce, Italy, [8]Oncologica Medica, Casa Sollievo della Sofferenza, San Giovanni Rotondo, Italy, [9]Direzione Scientifica, I.R.C.C.S. Istituto Tumori "Giovanni Paolo II", Bari, Italy, [10]Dipartimento di Biomedicina Traslazionale e Neuroscienze (DiBraiN), Università degli Studi di Bari "Aldo Moro", Bari, Italy

Objectives: Endometrial carcinosarcoma is a rare, aggressive high-grade endometrial cancer, accounting for about 5% of all uterine cancers and 15% of deaths from uterine cancers. The treatment can be complex, and the prognosis is poor. Its increasing incidence underscores the urgent requirement for personalized approaches in managing such challenging diseases.

Method: In this work, we designed an explainable machine learning approach to predict recurrence-free survival in patients affected by endometrial carcinosarcoma. For this purpose, we exploited the predictive power of clinical and histopathological data, as well as chemotherapy and surgical information collected for a cohort of 80 patients monitored over time. Among these patients, 32.5% have experienced the appearance of a recurrence.

Results: The designed model was able to well describe the observed sequence of events, providing a reliable ranking of the survival times based on the individual risk scores, and achieving a C-index equals to 70.00% (95% CI, 59.38−84.74).

Conclusion: Accordingly, machine learning methods could support clinicians in discriminating between endometrial carcinosarcoma patients at low-risk or high-risk of recurrence, in a non-invasive and inexpensive way. To the best of our knowledge, this is the first study proposing a preliminary approach addressing this task.

KEYWORDS

endometrial carcinosarcoma, recurrence-free survival, machine learning, explainable artificial intelligence, personalized medicine

Introduction

Endometrial carcinosarcoma is a biphasic tumor with both carcinomatous (epithelial) and sarcomatous (mesenchymal) elements (Raffone et al., 2022). It is a rare, aggressive high-grade endometrial cancer, accounting for about 5% of all uterine cancers and nearly 20% of non-endometrioid endometrial carcinomas (Siegel et al., 2022). Although non-endometrioid tumors make up 10–20% of endometrial malignancies, they are responsible for over 40% of endometrial cancer deaths (Lu and Broaddus, 2020).

The treatment can be complex, including the need to perform surgery, platinum-based chemotherapy, and radiotherapy. Despite this, the prognosis remains poor (Travaglino et al., 2022). Median overall survival is less than 2 years, and the 5-year overall survival is under 30% (approximately 50% for early stage and 20% for advanced stage disease). Even patients with early-stage disease have a 45% 5-year recurrence rate and 50% 5-year disease-specific mortality (Toboni et al., 2021). The rising incidence and poor outcomes of endometrial carcinosarcoma highlight an unmet need of personalising the management of these challenging patients, in order to allow more informed and targeted decision-making even in the presence of a complex prognosis (Grasso et al., 2017).

Over the past few years, due to the increased availability of data and the greatest computing power, Artificial Intelligence (AI) has emerged as a potential tool to deal with this big data, with the aim of optimizing cancer research, improving clinical practice, and promoting precision in healthcare (Farina et al., 2022). Specifically, Machine Learning (ML) is a subfield of AI which exploits mathematical and statistical approaches to develop learning models able to detect hidden patterns in the data and to improve model performance (Cuocolo et al., 2020). Anyway, ML models are often related to the concept of "black box" whereby even the operators who engineered the model cannot explain the reasoning behind their predictions. Hence, eXplainable Artificial Intelligence (XAI) algorithms have been designed to enable users to understand and appropriately trust ML model decisions (Gunning et al., 2019; Doshi-Velez and Kim, 2017).

Recently, ML methods have been also employed in the identification of prognostic factors capable of predicting cancer survival and the risk of disease recurrence, with the purpose of supporting clinicians in optimizing the clinical follow-up plan (Wang et al., 2019). The ability of ML methods to handle survival data by modelling the relationship between the event of interest and the predictor variables, could allow the design of personalized therapeutic options, showing better accuracy than conventional statistical approaches (Mazzaki et al., 2021). As a matter of fact, ML algorithms are able to capture and shape non-linearities between variables without having to specify the form of the relationship between them *a priori*.

So far, in the state-of-the-art, several ML models able to predict early diagnosis, as well as response to therapy and disease recurrence in the gynaecological cancers field have been proposed (Fiste et al., 2022; Sheehy et al., 2023; Akazawa and Hashimoto, 2021). However, there is a lack of ML models to predict recurrence-free survival in patients affected by endometrial carcinosarcoma. To the best of our knowledge, in this study, we proposed the first explainable ML approach which addresses this task exploiting the predictive power of clinical and histopathological data, as well as chemotherapy and surgical information of 80 endometrial carcinosarcoma patients.

Materials and methods

Experimental dataset

From 1988 to 2021, a total of 80 female patients affected by endometrial carcinosarcoma were enrolled and monitored over time, with the purpose of supervising their clinical pathway and the prospective occurrence of a recurrence event. While monitoring, 26 patients (32.5%) have experienced the appearance of a recurrence, and 54 patients (67.5%) have not recurred.

For each patient, clinical and histopathological data, as well as chemotherapy and surgical information, were collected from the patients' medical records. A total of 11 features were compiled, comprising the occurrence of a recurrent event (abbr. Recurrence, values: yes, no), age at diagnosis, tumor stage (abbr. Stage, values: I-II, III-IV), tumor hystotype (abbr. Type, values: homologous MMMT, heterologous MMMT), tumor size (abbr. size, values: ≤4cm, >4 cm), type of surgery (abbr. surgery, values: laparoscopic-LPS, laparotomic-LPT), having performed the omentectomy (abbr. omentectomy, values: yes, no), having performed the lymphadenectomy (abbr. lymphadenectomy, values: yes, no), chemotherapy scheme received (abbr. CT scheme, values: CBDCA, no CBDCA). The observation time, intended as the time in months between the date of diagnosis and either the appearance of a recurrence for recurrent patients or the last follow-up for non-recurrent patients, was also recorded for each patient.

An overview about the sample clinical properties is provided by Table 1.

Study design

We performed a ML recurrence-free survival analysis to estimate the time it takes for recurrence events to occur depending on the combination of values assumed by features.

The two most important notions on which survival analysis is based are the survival and hazard functions (Kartsonaki, 2016). The *Survival function* is defined as the probability of an event of interest T occurring after a specified time t:

$$S(t) = \Pr(T > t)$$

On the other hand, the *Hazard function* represents the likelihood for an individual to experience the event in a short interval of time $t + \Delta t$, given that the event has not occurred before time t:

$$h(t) = \frac{\Pr(t < T \leq t + \Delta t | T > t)}{\Delta t}$$

A starting point for the analysis needs to be also defined, and the beginning of each subject's observation time coincides with this

TABLE 1 Clinical features distribution over the study population.

Feature	Distribution
Overall	80; 100%
Recurrence	
Yes (abs; %)	26; 32.5%
No (abs; %)	54; 67.5%
Age at diagnosis	
Median; [q_1, q_3]	67; [60, 74]
Stage	
I-II (abs; %)	45; 56.8%
III-IV (abs; %)	35; 43.8%
Type	
Homologous MMMT (abs; %)	33; 41.2%
Heterologous MMMT (abs; %)	47; 58.8%
Size	
≤ 4 cm (abs; %)	29; 36.2%
> 4 cm (abs; %)	51; 63.8%
Surgery	
LPS (abs; %)	4; 5%
LPT (abs; %)	76; 95%
Omentectomy	
Yes (abs; %)	23; 28.8%
No (abs; %)	57; 71.2%
Myometrial invasion	
Yes (abs; %)	72; 90%
No (abs; %)	8; 10%
Lymphadenectomy	
Yes (abs; %)	42; 52.5%
No (abs; %)	38; 47.5%
CT scheme	
CBDCA (abs; %)	60; 75%
No CBDCA (abs; %)	20; 25%
Observation time (mm)	
Median; [q_1, q_3]	24; [7.5, 72.5]

starting point, time at which all subjects have the same risk equals to zero of the event occurring.

In this study, we implemented a stratified 5-fold cross-validation scheme over 5 rounds starting from clinical data of all patients enrolled, by dividing the population into strata, so that the right number of cases are sampled from each stratum to guarantee that the test set is representative of the entire population.

The first step consisted in adopting a feature selection approach based on a recursive feature elimination technique to identify a subset of relevant features for the outcome prediction. Starting from the original set of features, this technique allows to identify the optimal subset by recursively eliminating less important features by means of a linear regression. Features are consequently ranked according to

their estimated significance, and only the most important ones are employed for further steps (Ambusaidi et al., 2016).

Then, we trained a Gradient Boosting (GB) algorithm to determine how the hazard function varied according to the associated features previously selected. The GB algorithm is a non-parametric supervised learning belonging to the category of ensemble methods: it sequentially combines the predictions of multiple simple models named *base learners*, allowing each new learner to correct the previous one and, consequently, to reinforce the overall model. This process enables the minimization of a specific loss function using a well-defined base learner (Nguyen, 2019). In this work, we optimized the partial likelihood loss of the Cox's proportional hazards model by means of a regression tree base learner (Kleinbaum and Klein, 2012; Breiman et al., 2017). Specifically, a total of 100 estimators were employed. The model was trained setting the other hyperparameters as default (Pölsterl, 2020).

The discrimination power of the above-mentioned model was then evaluated in terms of the Concordance-index (C-index), that represents the model's ability to correctly provide a reliable ranking of the survival times based on the individual risk scores (Longato et al., 2020).

According to GB model predictions, for each patient in the test set both a predicted survival and a predicted cumulative hazard function were estimated and depicted. These are two stepwise functions in which the occurrence of one or more events is represented by a vertical drop or slope, respectively.

Finally, we adopted a XAI algorithm named Surv Local Interpretable Model-agnostic Explanations (SurvLIME), to explain the contribution of each of the most significant features to the decision of the ML survival model, both at patient and at all dataset levels. In both cases, this method allows to compute local interpretability by providing a ranking among the set of features, even considering the time space to give explanations with the goal of detecting possible dependencies between the features and the time (Kovalev et al., 2020). Particularly, to make these final predictions and explanations, we identified the most performing model among all models trained within the 5-fold cross-validation scheme, and we considered the same test set preliminarily identified according to the cross-validation framework. Therefore, only a subset was considered in this process to avoid overly optimistic estimates.

The idea behind the SurvLIME algorithm is to approximate the output produced by the ML survival model which has to be explained, by the output produced by a model belonging to a set of explanation models. Specifically, this approximation model is trained on new perturbed samples generated with the corresponding predictions of the ML survival model, by solving an optimization problem which minimizes the distance between the explanation and the prediction of the ML survival model. The approximation model adopted by the SurvLIME algorithm is the Cox proportional Hazards model, that is a semi-parametric survival algorithm whose output is the result of a multiple regression (Kleinbaum and Klein, 2012).

All the analysis steps have been performed by using Python.

Results

At the end of each cross-validation round, features were ranked in descending order according to their estimated significance, and only

features with a rank ≤3 were selected. The frequency of selection of each feature within all rounds of the cross-validation procedure is shown in Figure 1. Four variables, namely, myometrial invasion, omentectomy, surgery and histotype were always selected, presenting a frequency equals to 100%. Conversely, the age at diagnosis has never been selected as an important feature over the training process.

The designed ML survival model, trained on features resulted important by turns, was able to well describe the observed sequence of events, and its discrimination power was evaluated in terms of C-index along with its 95% confidence interval: 70.00% [59.38–84.74].

The model predictions also allowed to describe both a survival and a cumulative hazard function for each patient. The respective functions estimated by means of the best performing model are depicted in Figure 2. Due to the specularity of these functions, in both cases the curves resulted well separated into two groups. The first group of patients was characterized by a lower survival probability and, consequently, a higher risk of recurrence since the first months after diagnosis. Conversely, patients belonging to the second group were identified by a survival probability always greater than 80%, and which trend remained constant even after several months after diagnosis. A comparison highlighted that

patients with a higher risk of recurrence all share the following feature values: a heterologous MMMT type, a CBDCA CT scheme and a laparotomic-LPT surgery performed.
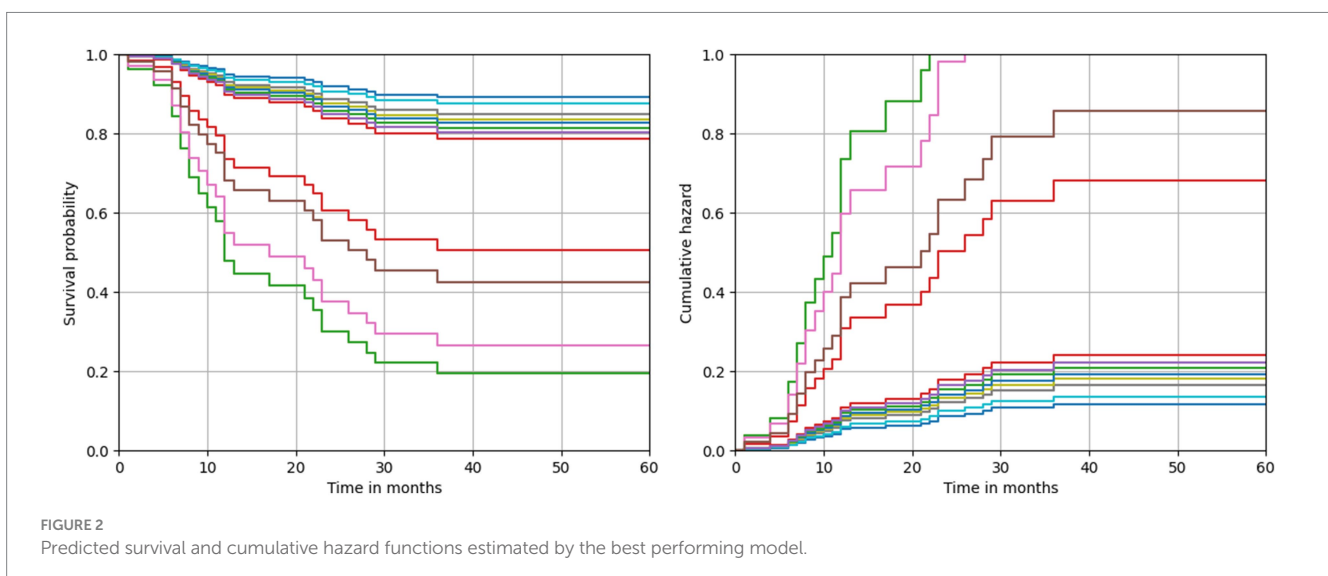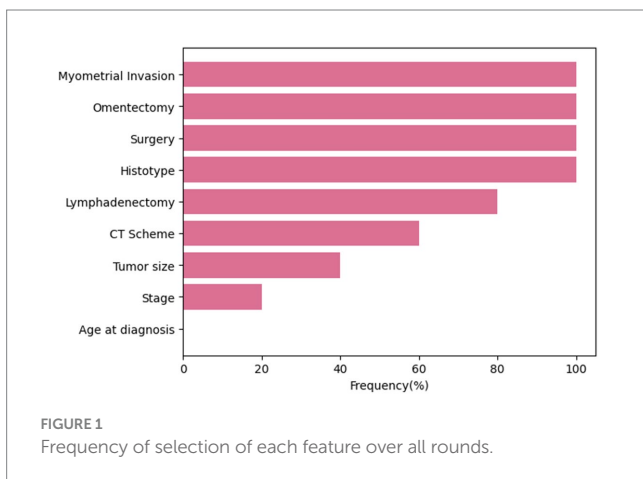
Despite the good performances of our ML survival model in predicting recurrence-free survival, the reasoning behind its predictions is unknown. To this aim, we provided local explanations of the contribution of important features to the model prediction, both at patient and all dataset level. Figures 3, 4 shows some examples of explanations at patient level. Each explanation consists of a feature importance diagram in which the feature contributions to the outcome are displayed in descending order, using a red colour palette for the features that increase the Cumulative Hazard Function and a blue palette for those that decrease it.

Considering Figure 3, for both patients the ML survival model correctly returned a high predicted hazard. This can be related to having performed the lymphadenectomy, a CBDCA CT scheme and a laparotomic-LPT surgery. On the other hand, patients illustrated in Figure 4 are related to a correctly low predicted hazard. In these cases, explanation diagrams highlighted how differences among their feature values are related to different feature contributions in terms of weight to the outcome prediction.

Lastly, Figure 5 depicts the local explanation at all dataset level. In this feature importance diagram, feature contributions are displayed in descending order by means of box plots representing the feature contribution distributions computed over the entire dataset. The feature positive or negative contribution in increasing the Cumulative Hazard Function is pictured by a red colour or blue colour palette, respectively.

Discussion

Uterine carcinosarcoma is a high-grade tumor including both epithelial and mesenchymal malignant cell components. Typically, the former shows low differentiation and a mix of characteristics, possibly displaying traits like endometrioid, clear-cell, or serous features. Tumor cells may organize in gland-like structures. The latter can resemble either endometrial stromal sarcoma or
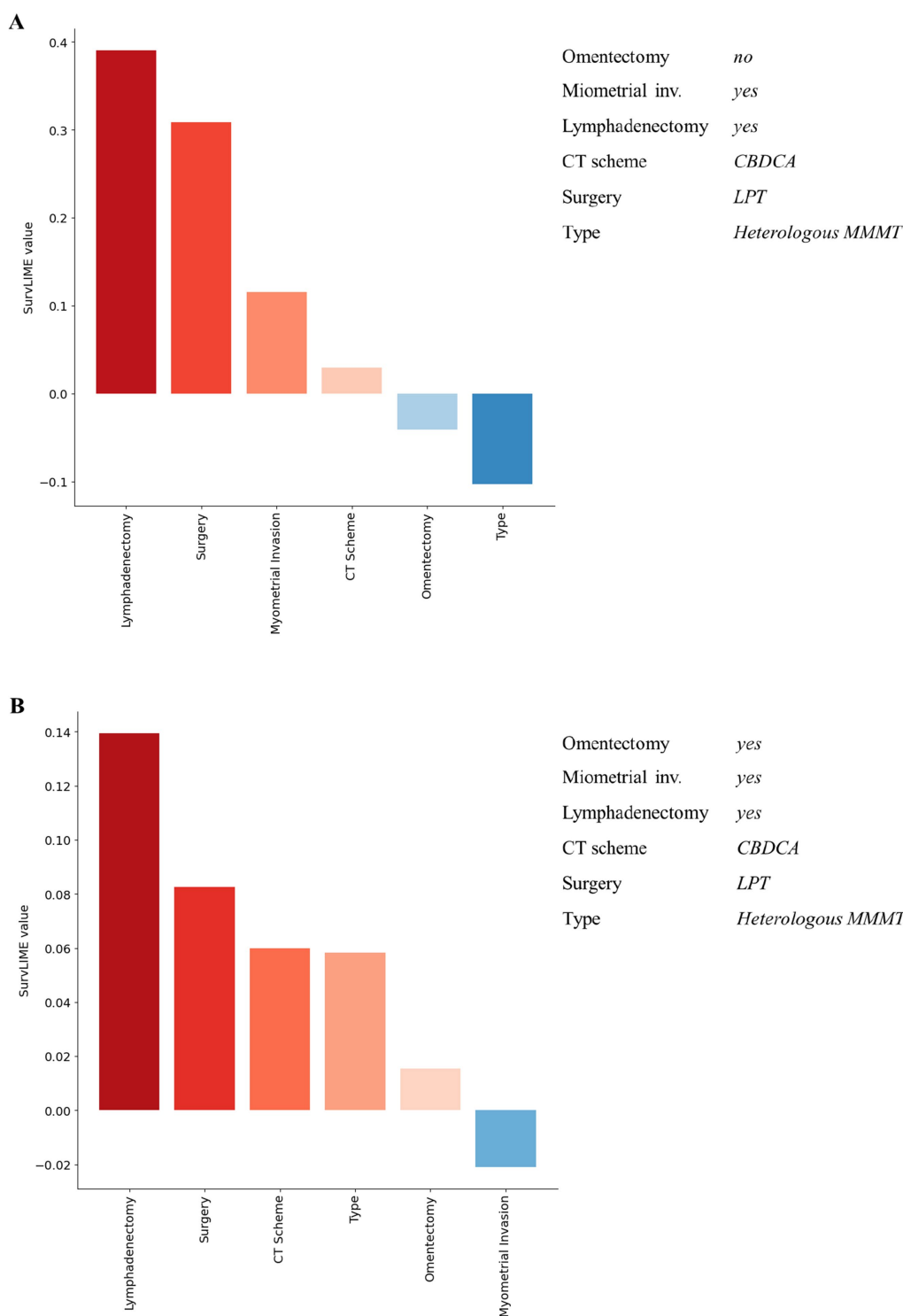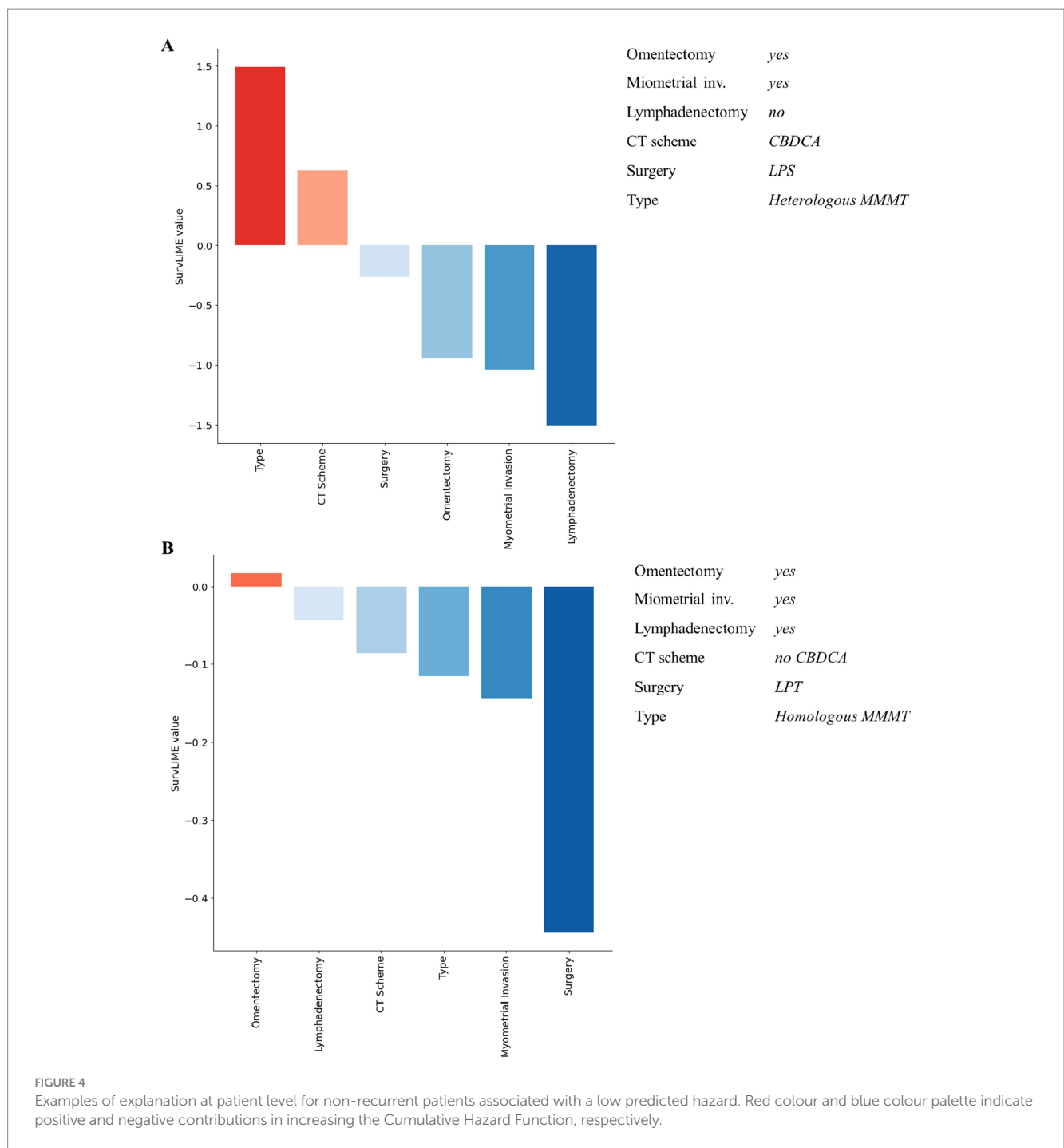

FIGURE 1
Frequency of selection of each feature over all rounds.

FIGURE 2
Predicted survival and cumulative hazard functions estimated by the best performing model.

A

Omentectomy	no
Miometrial inv.	yes
Lymphadenectomy	yes
CT scheme	CBDCA
Surgery	LPT
Type	Heterologous MMMT

B

Omentectomy	yes
Miometrial inv.	yes
Lymphadenectomy	yes
CT scheme	CBDCA
Surgery	LPT
Type	Heterologous MMMT

FIGURE 3
Examples of explanation at patient level for recurrent patients associated with a high predicted hazard. Red colour and blue colour palette indicate positive and negative contributions for increasing the Cumulative Hazard Function, respectively.

leiomyosarcoma, known as "homologous," or it may exhibit features akin to specialized connective tissues outside the uterus, such as muscle, cartilage, and bone, termed "heterologous." In both scenarios, angiolymphatic invasion is frequently observed (Bogani et al., 2023).
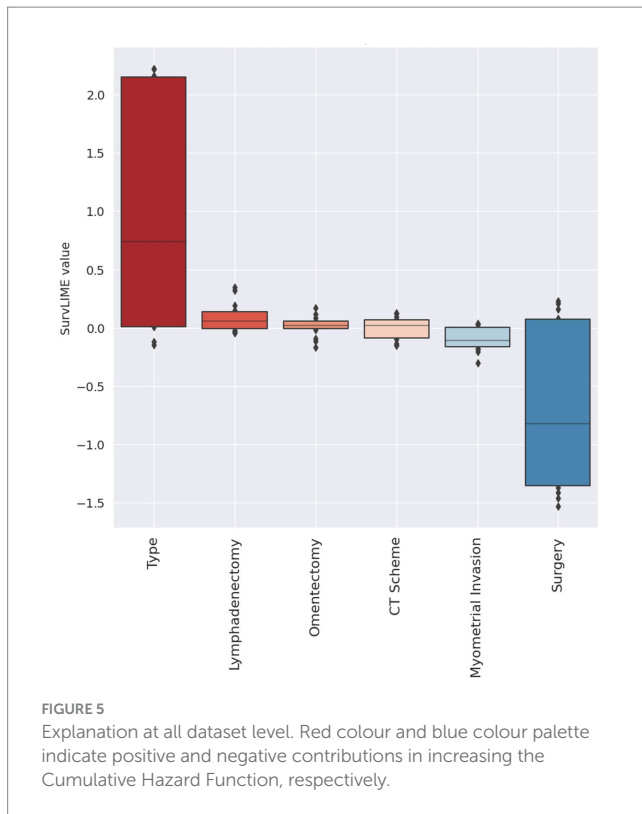
Despite surgical treatment and timely adjuvant multimodal therapy, more than half of the cases of endometrial carcinosarcoma will recur within the first 2 years (Concin et al., 2021). The management of the recurrent disease is highly personalized and should consider several factors, such as the performance status of the

FIGURE 4
Examples of explanation at patient level for non-recurrent patients associated with a low predicted hazard. Red colour and blue colour palette indicate positive and negative contributions in increasing the Cumulative Hazard Function, respectively.

patient, the size and sites of recurrences, and prior therapies (Pezzicoli et al., 2021). Importantly, it depends on whether the relapse is locoregional, oligometastatic, or disseminated and, second, on whether the patient has already received radiotherapy, as radiotherapy rechallenge is generally avoided for safety reasons. Again, the best treatment approach is multimodal. Patients with recurrent disease (including peritoneal and lymph node relapse) should be considered for surgery only if it is anticipated that complete removal of macroscopic disease can be achieved with acceptable morbidity and be treated in specialized centres (Beckmann et al., 2021). External beam radiotherapy can be used in radiotherapy-naïve patients or those who had received only prior vaginal brachytherapy.

Immunotherapy (with or without tyrosine kinase inhibitor) is the emerging preferred second-line systemic treatment. After the failure of immunotherapy, chemotherapy alone (generally mono-chemotherapy) is the preferred treatment in cases of disseminated metastases (Abu-Rustum et al., 2021).

Owing to the rare and aggressive nature of endometrial carcinosarcoma, the complexity of its management both at diagnosis and recurrence, as well as its high recurrence rate (Amant et al., 2005), identifying an ensemble of prognostic factors able to accurately predict recurrence-free survival in patients affected by this malignancy could allow more informed and targeted decision-making, enabling proactive clinical management

FIGURE 5
Explanation at all dataset level. Red colour and blue colour palette indicate positive and negative contributions in increasing the Cumulative Hazard Function, respectively.

even in the presence of a complex prognosis. Actually, the ability to identify patients at a major risk of recurrence at an early stage could allow clinicians to tailor treatment plans, both adopting more aggressive strategies such as intensive or combination chemotherapy protocols, adjuvant radiotherapy or experimental approaches in patients with a high probability of recurrence and sparing low-risk patients from unnecessary interventions. Moreover, an accurate predictive model can guide the frequency and intensity of clinical follow-up, allowing an early detection of recurrence and improving the likelihood of disease control. Finally, thanks to predictive modelling, it is possible to more selectively enrol patients which may be candidates for clinical trials of new drugs or experimental therapies, especially when standard options have obvious limitations.

Over the past years, several efforts have been made to develop a greater awareness and deeper understanding of endometrial carcinosarcoma pathogenesis, with the purpose of identifying new targeted therapies and providing specific guidelines for the management of this tumor (Bogani et al., 2023). Besides, endometrial cancer treatment has provided new changes by incorporating biological, clinical, genomic, and clinico-pathologic characteristics of the women affected by this tumor, and recent studies showed that molecular targets such as L1CAM (L1 cell adhesion molecule) plays an important role as prognostic factor and could provide a potential useful tool for tailoring the need of adjuvant therapy (Giannini et al., 2024; Vizza et al., 2021). As well, a prognostic nomogram to predict the overall survival rate in endometrial carcinosarcoma patients by exploiting lymph-node metastasis information has been proposed (Gao et al., 2021).

However, there is a lack of research studies focused on the prediction of disease recurrence risk.

In this study, we proposed the first explainable ML method designed to predict recurrence-free survival in patients affected by endometrial carcinosarcoma. The nested feature importance approach allowed us to identify the most relevant variables for this outcome prediction. Accordingly, promising results were achieved in providing a reliable ranking of the survival times based on the individual risk scores (C-index: 70%). Finally, with the aim to enable clinicians to understand the reasoning behind the ML model predictions, the implemented XAI algorithm computed the contribution of each of the most significant features to the model decision, both at patient and at all dataset levels.

To conclude, the proposed explainable ML model represents the first effort in devising an artificial intelligence-based tool to be enclosed in clinical practice to support clinicians in discriminating between endometrial carcinosarcoma patients at low-risk or high-risk of recurrence in a non-invasive and inexpensive way, also providing an intelligible explanation on how the clinical characteristics considered for those patients contributed to the estimated risk. Accordingly, the ability of this model in detecting the risk for a patient of experiencing recurrence could aid clinicians to personalise therapeutic options, by candidating high-risk patients to adjuvant chemotherapy and saving low-risk patients from unnecessary aggressive treatments.

Besides, a limitation of our study deals with its retrospective design and the limited dimension of the dataset. As far as the limited dataset size, this could affect the robustness and generalizability of the ML model which generally require larger dataset, in contrast to what classic survival approaches, such as Cox regression, need. However, the advantages in exploiting ML techniques rather than classical methods are the increased flexibility, the ability to adapt to non-linear relationships, together with improved predictive performances. Actually, relationships between variables are often non-linear or complex, and some effects may depend on interactions between them. ML algorithms are able to capture and shape these non-linearities and interactions without having to specify the form of the relationship between variables *a priori*. Definitely, employing ML models with larger datasets, it could be possible to achieve higher performances and improve the model. For this purpose, in our future work we will collect an external dataset for prospective validation, in order to establish documented evidence that the model is able to consistently produce the desired results within predetermined specifications and quality attributes.

Data availability statement

The data analyzed in this study is subject to the following licenses/restrictions: data from this study are available upon request since data contain potentially sensitive information. The data request may be sent to the scientific direction (e-mail: dirscientifica@oncologico.bari.it).

Ethics statement

The studies involving humans were approved by Institutional Ethics Committee—IRCCS Istituto Tumori Giovanni Paolo II, Bari.

The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

SB: Conceptualization, Data curation, Formal analysis, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. FA: Conceptualization, Data curation, Formal analysis, Methodology, Validation, Writing – original draft, Writing – review & editing. GCo: Conceptualization, Data curation, Formal analysis, Resources, Supervision, Writing – original draft, Writing – review & editing. ES: Data curation, Writing – review & editing. AC: Data curation, Writing – review & editing. MC: Data curation, Formal analysis, Writing – original draft, Writing – review & editing. AF: Data curation, Formal analysis, Writing – original draft, Writing – review & editing. GA: Writing – review & editing. GCa: Writing – review & editing. GD: Writing – review & editing. BM: Data curation, Writing – review & editing. EN: Data curation, Writing – review & editing. AL: Writing – review & editing. EV: Writing – review & editing. VL: Writing – review & editing, Conceptualization, Data curation, Formal analysis, Resources, Supervision, Writing – original draft.

RM: Conceptualization, Data curation, Formal analysis, Resources, Supervision, Writing – original draft, Writing – review & editing.

Conflict of interest

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

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Abu-Rustum, N. R., Yashar, C. M., Bradley, K., Campos, S. M., Chino, J., Chon, H. S., et al. (2021). NCCN guidelines® insights: uterine neoplasms, version 3.2021. *J. Natl. Compr. Cancer Netw.* 19, 888–895. doi: 10.6004/jnccn.2021.0038

Akazawa, M., and Hashimoto, K. (2021). Artificial intelligence in gynecologic cancers: current status and future challenges – a systematic review. *Artif. Intell. Med.* 120:102164. doi: 10.1016/j.artmed.2021.102164

Amant, F., Moerman, P., Neven, P., Timmerman, D., van Limbergen, E., and Vergote, I. (2005). Endometrial cancer. *Lancet* 366, 491–505. doi: 10.1016/S0140-6736(05)67063-8

Ambusaidi, M. A., He, X., Nanda, P., and Tan, Z. (2016). Building an intrusion detection system using a filter-based feature selection algorithm. *IEEE Trans. Comput.* 65, 2986–2998. doi: 10.1109/TC.2016.2519914

Beckmann, K., Selva-Nayagam, S., Olver, I., Miller, C., Buckley, E. S., Powell, K., et al. (2021). Carcinosarcomas of the uterus: prognostic factors and impact of adjuvant treatment. *Cancer Manag. Res.* 13, 4633–4645. doi: 10.2147/CMAR.S309551

Bogani, G., Ray-Coquard, I., Concin, N., Ngoi, N. Y. L., Morice, P., Caruso, G., et al. (2023). Endometrial carcinosarcoma. *Int. J. Gynecol. Cancer* 33, 147–174. doi: 10.1136/ijgc-2022-004073

Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J. (2017). Classification and regression trees. Chapman and Hall/CRC: Routledge. doi: 10.1201/9781315139470

Concin, N., Matias-Guiu, X., Vergote, I., Cibula, D., Mirza, M. R., Marnitz, S., et al. (2021). ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. *Int. J. Gynecol. Cancer* 31, 12–39. doi: 10.1136/ijgc-2020-002230

Cuocolo, R., Caruso, M., Perillo, T., Ugga, L., and Petretta, M. (2020). Machine learning in oncology: a clinical appraisal. *Cancer Lett.* 481, 55–62. doi: 10.1016/j.canlet.2020.03.032

Doshi-Velez, F, and Kim, B (2017) Towards a rigorous science of interpretable machine learning. doi: 10.48550/arXiv.1702.08608

Farina, E., Nabhen, J. J., Dacoregio, M. I., Batalini, F., and Moraes, F. Y. (2022). An overview of artificial intelligence in oncology. *Future Sci. OA* 8:FSO787. doi: 10.2144/fsoa-2021-0074

Fiste, O., Liontos, M., Zagouri, F., Stamatakos, G., and Dimopoulos, M. A. (2022). Machine learning applications in gynecological cancer: a critical review. *Crit. Rev. Oncol. Hematol.* 179:103808. doi: 10.1016/j.critrevonc.2022.103808

Gao, L., Lyu, J., Luo, X., Zhang, D., Jiang, G., Zhang, X., et al. (2021). Nomogram to predict overall survival based on the log odds of positive lymph nodes for patients with endometrial carcinosarcoma after surgery. *BMC Cancer* 21:1149. doi: 10.1186/s12885-021-08888-0

Giannini, A., D'Oria, O., Corrado, G., Bruno, V., Sperduti, I., Bogani, G., et al. (2024). The role of L1CAM as predictor of poor prognosis in stage I endometrial cancer: a systematic review and meta-analysis. *Arch. Gynecol. Obstet.* 309, 789–799. doi: 10.1007/s00404-023-07149-8

Grasso, S., Loizzi, V., Minicucci, V., Resta, L., Camporeale, A. L., Cicinelli, E., et al. (2017). Malignant mixed Müllerian tumour of the uterus: analysis of 44 cases. *Oncology* 92, 197–204. doi: 10.1159/000452277

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S., and Yang, G. Z. (2019). XAI—Explainable artificial intelligence. *Sci. Robot.* 4:aay7120. doi: 10.1126/scirobotics.aay7120

Kartsonaki, C. (2016). Survival analysis. *Diagn. Histopathol.* 22, 263–270. doi: 10.1016/j.mpdhp.2016.06.005

Kleinbaum, D. G., and Klein, M. (2012) The cox proportional hazards model and its characteristics. In: *Survival analysis. Statistics for Biology and Health.* ed. (New York, NY: Springer). 159: 97. doi: 10.1007/978-1-4419-6646-9_3

Kovalev, MS, and Utkin, L V., Kasimov, EM (2020) Surv LIME: A method for explaining machine learning survival models, *knowledge-based systems.* 203:106164. doi: 10.1016/j.knosys.2020.106164

Longato, E., Vettoretti, M., and Di Camillo, B. (2020). A practical perspective on the concordance index for the evaluation and selection of prognostic time-to-event models. *J. Biomed. Inform.* 108:103496. doi: 10.1016/j.jbi.2020.103496

Lu, K. H., and Broaddus, R. R. (2020). Endometrial Cancer. *N. Engl. J. Med.* 383, 2053–2064. doi: 10.1056/NEJMra1514010

Mazzaki, J., Katsumata, K., Ohno, Y., Udo, R., Tago, T., Kasahara, K., et al. (2021). A novel prediction model for Colon Cancer recurrence using auto-artificial intelligence. *Anticancer Res.* 41, 4629–4636. doi: 10.21873/anticanres.15276

Nguyen, NP (2019) Gradient boosting for survival analysis with applications in oncology. *USF Tampa Graduate Theses and Dissertations.* Available at: https://digitalcommons.usf.edu/etd/8062

Pezzicoli, G., Moscaritolo, F., Silvestris, E., Silvestris, F., Cormio, G., Porta, C., et al. (2021). Uterine carcinosarcoma: An overview. *Crit. Rev. Oncol. Hematol.* 163:103369:103369. doi: 10.1016/j.critrevonc.2021.103369

Pölsterl, S. (2020). Scikit-survival: a library for time-to-event analysis built on top of scikit-learn. *J. Mach. Learn. Res.* 21, 1–6.

Raffone, A., Travaglino, A., Raimondo, D., Maletta, M., de Vivo, V., Visiello, U., et al. (2022). Uterine carcinosarcoma vs endometrial serous and clear cell carcinoma: a systematic review and meta-analysis of survival. *Int. J. Gynecol. Obstet.* 158, 520–527. doi: 10.1002/ijgo.14033

Sheehy, J., Rutledge, H., Acharya, U. R., Loh, H. W., Gururajan, R., Tao, X., et al. (2023). Gynecological cancer prognosis using machine learning techniques: a systematic review of the last three decades (1990–2022). *Artif. Intell. Med.* 139:102536:102536. doi: 10.1016/j.artmed.2023.102536

Siegel, R. L., Miller, K. D., Fuchs, H. E., and Jemal, A. (2022). Cancer statistics, 2022. *CA Cancer J. Clin.* 72, 7–33. doi: 10.3322/caac.21708

Toboni, M. D., Crane, E. K., Brown, J., Shushkevich, A., Chiang, S., Slomovitz, B. M., et al. (2021). Uterine carcinosarcomas: from pathology to practice. *Gynecol. Oncol.* 162, 235–241. doi: 10.1016/j.ygyno.2021.05.003

Travaglino, A., Raffone, A., Raimondo, D., Arciuolo, D., Angelico, G., Valente, M., et al. (2022). Prognostic value of the TCGA molecular classification in uterine carcinosarcoma. *Int. J. Gynecol. Obstet.* 158, 13–20. doi: 10.1002/ijgo.13937

Vizza, E., Bruno, V., Cutillo, G., Mancini, E., Sperduti, I., Patrizi, L., et al. (2021). Prognostic role of the removed vaginal cuff and its correlation with L1CAM in low-risk endometrial adenocarcinoma. *Cancers (Basel)* 14:34. doi: 10.3390/cancers14010034

Wang, P., Li, Y., and Reddy, C. K. (2019). Machine learning for survival analysis. *ACM Comput. Surv.* 51, 1–36. doi: 10.1145/3214306