Artificial Intelligence In Medicine 139 (2023) 102536

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed

Gynecological cancer prognosis using machine learning techniques: A systematic review of the last three decades (1990–2022)

Joshua Sheehy [a], Hamish Rutledge [a], U. Rajendra Acharya [b], Hui Wen Loh [c], Raj Gururajan [d], Xiaohui Tao [e], Xujuan Zhou [e], Yuefeng Li [f], Tiana Gurney [a], Srinivas Kondalsamy-Chennakesavan [a,*]

[a] The University of Queensland Rural Clinical School, Locked Bag 9009, Toowoomba DC 4350, Queensland, Australia
[b] School of Engineering, Ngee Ann Polytechnic, Singapore 599489, Singapore
[c] School of Science and Technology, Singapore University of Social Sciences, Singapore 599494, Singapore
[d] University of Southern Queensland, 37 Sinnathamby Blvd, Springfield Central, Qld 4300, Australia
[e] University of Southern Queensland, 487 – 535 West Street, Toowoomba, Qld 4350, Australia
[f] School of Computer Science, Queensland University of Technology, Brisbane, Australia

ARTICLE INFO

Keywords:
Machine learning
Artificial intelligence
Gynecological oncology
Prognosis
Prediction

ABSTRACT

Objective: Many Computer Aided Prognostic (CAP) systems based on machine learning techniques have been proposed in the field of oncology. The objective of this systematic review was to assess and critically appraise the methodologies and approaches used in predicting the prognosis of gynecological cancers using CAPs.
Methods: Electronic databases were used to systematically search for studies utilizing machine learning methods in gynecological cancers. Study risk of bias (ROB) and applicability were assessed using the PROBAST tool. 139 studies met the inclusion criteria, of which 71 predicted outcomes for ovarian cancer patients, 41 predicted outcomes for cervical cancer patients, 28 predicted outcomes for uterine cancer patients, and 2 predicted outcomes for gynecological malignancies broadly.
Results: Random forest (22.30 %) and support vector machine (21.58 %) classifiers were used most commonly. Use of clinicopathological, genomic and radiomic data as predictors was observed in 48.20 %, 51.08 % and 17.27 % of studies, respectively, with some studies using multiple modalities. 21.58 % of studies were externally validated. Twenty-three individual studies compared ML and non-ML methods. Study quality was highly variable and methodologies, statistical reporting and outcome measures were inconsistent, preventing generalized commentary or meta-analysis of performance outcomes.
Conclusion: There is significant variability in model development when prognosticating gynecological malignancies with respect to variable selection, machine learning (ML) methods and endpoint selection. This heterogeneity prevents meta-analysis and conclusions regarding the superiority of ML methods. Furthermore, PROBAST-mediated ROB and applicability analysis demonstrates concern for the translatability of existing models. This review identifies ways that this can be improved upon in future works to develop robust, clinically translatable models within this promising field.

1. Introduction

The National Cancer Institute defines cancer as a "disease in which some of the body's cells grow uncontrollably and spread to other parts of the body" [1]. When such uncontrolled growth occurs in women's reproductive organs or genitals, they are referred to as 'gynecological cancers'. There are five main types of gynecological cancers (cervical, ovarian, uterine, vaginal and vulval), named after the organ or tissue from which they originate. The most common gynecological malignancies – ovarian, cervical and uterine cancers – present a significant disease burden worldwide [2]. These cancers are prognostically variable. Of these, ovarian cancer has the highest rate of recurrence, at 85 %

* Corresponding author.
E-mail addresses: joshua.sheehy@uq.net.au (J. Sheehy), hamish.rutledge@uq.net.au (H. Rutledge), rajendra_udyavara_acharya@np.edu.sg (U.R. Acharya), hwloh002@suss.edu.sg (H.W. Loh), Raj.Gururajan@usq.edu.au (R. Gururajan), Xiaohui.Tao@usq.edu.au (X. Tao), xujuan.zhou@usq.edu.au (X. Zhou), y2.li@qut.edu.au (Y. Li), t.gurney@uq.edu.au (T. Gurney), uqskonda@uq.edu.au (S. Kondalsamy-Chennakesavan).

https://doi.org/10.1016/j.artmed.2023.102536
Received 27 October 2021; Received in revised form 19 March 2023; Accepted 23 March 2023
Available online 29 March 2023

[3], and lowest rate of five-year survival at 30 % [4]. This is worsened by its non-specific symptomology and frequent late-stage diagnosis [5]. Although some clinical factors are prognostic, such as grade, stage, tumor subtype and debulking surgery success, the most commonly used clinical predictors [5] and biomarkers [6] are inadequate to predict clinical outcomes. Cervical cancer is one of the most common gynecological malignancies and is the fourth highest cause of cancer mortality in women worldwide [2]. Although some high Human Development Index (HDI) nations have had success in reducing the disease burden with screening and prevention programs [7], the prognostication of advanced-stage cervical cancer is variable [8]. Finally, although uterine cancer has a better prognosis than other malignancies, this disease is often very heterogeneous, making prognostication with current methods a challenge [9].

In recent years, prognostication has developed as a major focus in oncology, where decision making is influenced by the predicted probability of future events [10]. Treatment for gynecological cancers depends on the extent to which they have spread and the type of cancer, and includes modalities such as surgery, chemotherapy and radiotherapy. Developing oncological prediction algorithms and decision support tools would be useful for allowing clinicians to choose optimal screening, therapeutic and follow-up pathways for patients. However, challenges arise from the cancers' biological complexity and prognostic variability, alongside the ever-changing clinical, biological and pathological understanding of these malignancies [11].

Healthcare systems and clinicians currently use several tools to screen, diagnose and treat patients; however, current clinical approaches for many malignancies favor clinical staging and histopathological parameters with multivariate modelling showing limited success [12]. To address these shortcomings, machine learning (ML) approaches have been used to facilitate complex prognostic modelling that may outperform traditional methods [13]. ML methods aim to develop predictive algorithms without requiring complete prior rule definition, a valuable approach in complex clinical settings [12]. Predictive systems begin with data that undergoes pre-processing and feature extraction, followed by statistical analysis of extracted features, with selected features producing a classification result (Fig. 1). ML systems can be used at each step of this process. The ML classifiers (shown in Appendix A) used to perform the classification tasks include both unsupervised learning, which draws correlations within a dataset without a directed outcome, and supervised learning methods, including support vector machines (SVMs) and artificial neural networks (ANNs), which are goal-directed toward a particular outcome, regression or classification [12]. ML systems typically undergo training on a 'training' dataset and use a 'validation' dataset that the system is naïve to, to facilitate assessment of its performance, while still tuning its parameters. Finally, the system is typically exposed to a 'testing' dataset that it is naïve to, to facilitate an unbiased assessment of the final model's performance.

There is a paucity of clinical translatability of these methods for gynecological malignancies. Although they have been studied in this setting, they are variable in both approach and success. Systematic reviews have previously broadly summarized artificial intelligence (AI) in gynecologic imaging [14], or the application of ML methods broadly to gynecological cancers [15,16]. However, to the authors' knowledge, a systematic review of the literature specific to prognostication in gynecological cancer has yet to be performed. Therefore, this study aims to systematically review ML in the prognostication of gynecological malignancies and evaluate the methodologies used.

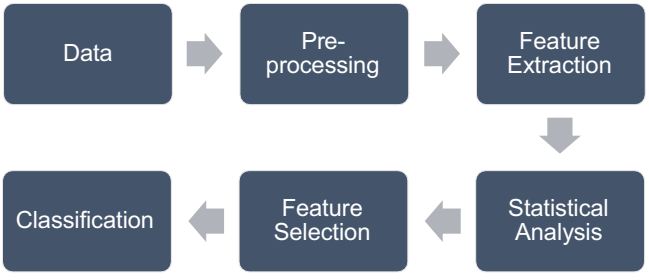

Fig. 1. Conceptual clinical predictive rule stepwise process. Clinical prediction systems begin with data that undergoes pre- processing and feature extraction, followed by statistical analysis and subsequent feature selection to produce a classification result. Machine learning processes can be used at each step of this process.

2. Materials and methods

2.1. Search strategy

This study was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [17]. We searched PubMed, Embase, Web of Science, ENGINE, Scopus, IEEE Xplore and ACM Digital Library for studies exploring the use of ML methods for predicting prognosis for gynecological malignancies. Our search query was developed in PubMed using MeSH and keyword terms, and then revised for the other databases (Supplementary Table 1). The final iteration of the search was completed on 01 September 2022 and included all results available up to the search date.

2.2. Study selection

We utilized a three-stage screening process to assess the suitability of articles identified by our database searches (Fig. 2). This screening process was completed by two authors (JS and HR). Results for which there was uncertainty were resolved by consensus. For those results where disagreement persisted, a third author (SKC) was consulted for adjudication. All articles were initially screened by title and subsequently by abstract. For articles identified as being potentially eligible, a full-text review was completed to confirm suitability.

2.3. Inclusion and exclusion criteria

Studies were accepted if they fulfilled the following inclusion criteria:

- A focus on predicting prognostic clinical outcomes after diagnosis in gynecological malignancy, using ML techniques
- Reporting primary research, and excluding reviews
- Sufficiently detailed reporting of methods and results for credible data extraction
- Available in English

2.4. Quality assessment

The PROBAST (Prediction model Risk Of Bias ASsessment Tool) is a ROB and clinical applicability assessment tool developed specifically for diagnostic and prognostic models [18,19]. This tool assesses ROB across four domains – participants, predictors, outcomes and analysis – using 20 signaling questions. ROB is assessed across all four domains and the study overall as low, unclear or high. Applicability is graded as low, unclear or high concern for three domains – participants, predictors and outcomes – and the study overall. This tool was applied to each study following recommendations made by its designing authors, with multiple authors cross-checking assessments to ensure consistency. PROBAST results are also presented here according to recommendations outlined by the developing authors.

2.5. Data extraction

Data extraction was completed using a structured proforma designed by three of the authors with relevant expertise (SKC, XT and XZ). Data

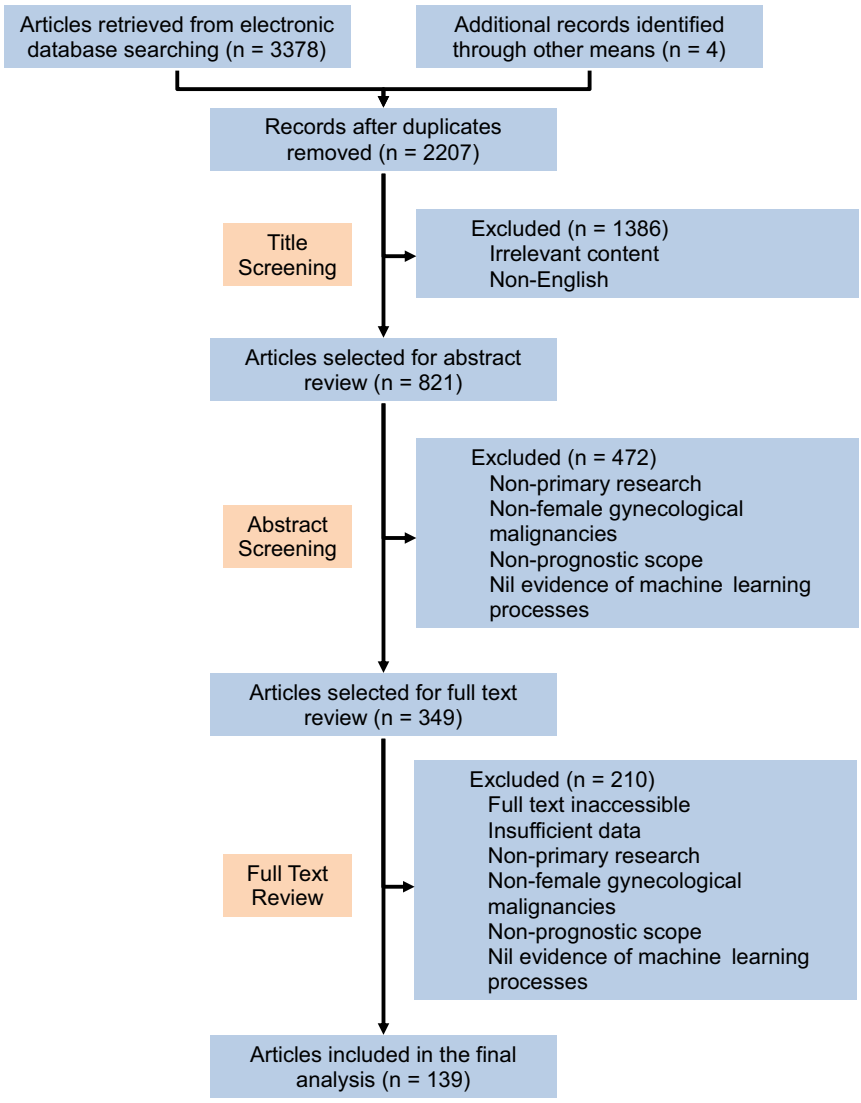

Fig. 2. PRISMA flow diagram for search strategy. Databases searched were PubMed, Embase, Web of Science, ENGINE, Scopus, IEEE Xplore and ACM Digital Library.

were collected pertaining to the following categories: study characteristics (authors, year of publication, study type, study purpose), study dataset (sample source, time period, demographics, clinical characteristics, sample size), cancer type, variables (category, number tested, number modelled, pre-processing for selection of modelled variables), ML methods utilized, measured outcomes, validation (internal and external validation with validation outcomes), analysis metrics and results. ML classifiers were coded to their category, and prediction variables used in each study were coded as radiomics, genomics or clinicopathological.

Data extraction for all accepted articles was completed by two authors (JS and HR), with oversight by another author (SKC) to resolve discrepancies. Tables were generated in Microsoft Excel; figures were generated in Microsoft Powerpoint, Microsoft Excel or GraphPad Prism (version 9).

3. Results

The aim of this study was to systematically review the use of ML in the prognostication of gynecological malignancies and evaluate the methodologies used. In total, the initial search yielded 2207 unique papers with 349 papers passing title and abstract screening. 139 papers met all criteria for inclusion in the study [20–158]. Of these, 71 predicted outcomes for ovarian cancer patients, 41 predicted outcomes for cervical cancer patients, 28 predicted outcomes for uterine cancer patients (including endometrial cancer) and two predicted outcomes for gynecological cancer generally. There were no studies that examined vulval or vaginal cancers that met the inclusion criteria. All included studies were published in peer reviewed journals, with publication dates ranging from 1993 to 2022.

Overall summary statistics for the included literature are presented in Table 1, while paper-specific data are presented in Supplementary Table 2. The volume of research done in this field has grown rapidly with time, with 58.99 % ($n = 82$) of papers published within the past three years at the time of the literature search. Most papers utilized only text-based inputs to their ML tools (92.09 %), with insufficient papers ($n = 10$) using image-based inputs to compare use of different input formats for different ML tools. Different data sources were used to assess different variables classes; hospital data were predominantly used to assess clinicopathological and radiomic variables (89.71 %), while data from databases were predominantly used to assess genomic variables (82.05 %). 65.47 % of studies had clear and appropriate inclusion or exclusion criteria, while the remainder did not. There was significant heterogeneity in the number of variables tested for modelling between different studies, ranging from 30 to 46,773 variables used in the included studies' final prediction model.

Table 1
Summary statistics for papers meeting inclusion criteria. Results provided for overall dataset as well as stratified by cancer type. Some studies utilized multiple ML methods in their final predictive tool; in these cases, each method was counted separately. Abbreviations used were as follows: ANN = Artificial Neural Network; BN = Bayesian Network; DL = Deep Learning; DT = Decision Tree; DFS = Disease-Free Survival; EL = Ensemble Learning; EN = Elastic Net; EV = External Validation; GBM = Gradient Boosting Machine; IQR = Interquartile Range; IV = Internal Validation; k-NN = k-Nearest Neighbor; MC = Markov Clustering; MK = Multiple Kernel; ML = Machine Learning; NB = Naïve Bayes; OS = Overall Survival; PFS = Progression-free Survival; RF = Random Forest; RFS = Recurrence-Free Survival; SVM = Support Vector Machine.

		Ovarian		Cervical		Uterine		Total	
Dataset type	Hospital	25	(35.21 %)	25	(60.98 %)	9	(32.14 %)	60	(43.17 %)
	Database	42	(59.15 %)	15	(36.59 %)	15	(53.57 %)	70	(50.36 %)
	Both	3	(4.23 %)	1	(2.44 %)	4	(14.29 %)	8	(5.76 %)
	Unclear	1	(1.41 %)	0	(0 %)	0	(0 %)	1	(0.72 %)
Sample size	Median	309.5		189		543		284.5	
	IQR	356.25		189		577.75		411	
Predictor variables	Clinicopathological	21	(29.58 %)	14	(34.15 %)	11	(39.29 %)	47	(33.81 %)
	Genomic	36	(50.7 %)	12	(29.27 %)	12	(42.86 %)	58	(41.73 %)
	Radiomic	2	(2.82 %)	11	(26.83 %)	0	(0 %)	13	(9.35 %)
	Clinicopathological/genomic	5	(7.04 %)	2	(4.88 %)	3	(10.71 %)	10	(7.19 %)
	Clinicopathological/radiomic	5	(7.04 %)	2	(4.88 %)	1	(3.57 %)	8	(5.76 %)
	Radiomic/genomic	0	(0 %)	0	(0 %)	1	(3.57 %)	1	(0.72 %)
	Clinicopathological/genomic/radiomic	2	(2.82 %)	0	(0 %)	0	(0 %)	2	(1.44 %)
ML methods	ANN	13	(18.31 %)	6	(14.63 %)	3	(10.71 %)	21	(15.11 %)
	RF	9	(12.68 %)	14	(34.15 %)	7	(25 %)	32	(23.02 %)
	SVM	19	(26.76 %)	8	(19.51 %)	3	(10.71 %)	30	(21.58 %)
	BN	0	(0 %)	0	(0 %)	2	(7.14 %)	2	(1.44 %)
	DL	2	(2.82 %)	2	(4.88 %)	0	(0 %)	4	(2.88 %)
	DT	9	(12.68 %)	1	(2.44 %)	3	(10.71 %)	12	(8.63 %)
	EL	3	(4.23 %)	1	(2.44 %)	1	(3.57 %)	5	(3.6 %)
	EN	0	(0 %)	4	(9.76 %)	2	(7.14 %)	6	(4.32 %)
	GALGO	1	(1.41 %)	0	(0 %)	0	(0 %)	1	(0.72 %)
	GBM	4	(5.63 %)	2	(4.88 %)	2	(7.14 %)	7	(5.04 %)
	ML Regression	9	(12.68 %)	6	(14.63 %)	4	(14.29 %)	19	(13.67 %)
	k-NN	0	(0 %)	1	(2.44 %)	1	(3.57 %)	2	(1.44 %)
	MC	2	(2.82 %)	0	(0 %)	0	(0 %)	2	(1.44 %)
	NB	1	(1.41 %)	1	(2.44 %)	0	(0 %)	2	(1.44 %)
	Unsupervised Clustering	4	(5.63 %)	0	(0 %)	1	(3.57 %)	5	(3.6 %)
	MK	1	(1.41 %)	0	(0 %)	0	(0 %)	1	(0.72 %)
Outcomes	OS only	38	(53.52 %)	18	(43.9 %)	17	(60.71 %)	72	(51.8 %)
	PFS only	4	(5.63 %)	1	(2.44 %)	1	(3.57 %)	5	(3.6 %)
	DFS only	2	(2.82 %)	2	(4.88 %)	0	(0 %)	4	(2.88 %)
	RFS only	17	(23.94 %)	11	(26.83 %)	5	(17.86 %)	34	(24.46 %)
	Combination	10	(14.08 %)	9	(21.95 %)	5	(17.86 %)	24	(17.27 %)
Validation	IV	18	(25.35 %)	7	(9.59 %)	5	(17.86 %)	30	(21.58 %)
	EV	48	(67.61 %)	33	(45.21 %)	22	(78.57 %)	101	(72.66 %)
	Not validated	5	(7.04 %)	1	(1.37 %)	1	(3.57 %)	8	(5.76 %)

3.1. Ovarian cancer

There were 71 papers reporting on ovarian cancer, with 25 studies (35.21 %) using hospital datasets, 42 (59.15 %) using database sets, three (4.23 %) using a combination of both, and one unclear. The median sample size was 309.5 with an interquartile range of 356.25. In terms of predictor variables, 43 papers (58.90 %) used genomic data; 36 papers (50.70 %) used genomic data only, five papers (7.04 %) combined genomic and clinicopathologic data, and two papers (2.82 %) combined genomic, clinicopathological and radiomic data. Thirty three papers (45.21 %) used clinicopathological data; 21 papers (29.57 %) used clinicopathological data only and five papers (7.04 %) each combined this with radiomic data or genomic data. Two studies (2.82 %) used radiomic data only.

In terms of ML methods, 19 studies (26.76 %) used SVMs, nine (12.67 %) used RFs, 13 (18.31 %) used ANNs, nine (12.68 %) used DTs and nine (12.68 %) used ML regression methods. Twenty three studies compared multiple ML methods, showing variable results, with no clear pattern of superiority when directly compared.

Eleven studies compared ML methods to non-ML prediction approaches, and all showed ML as superior [20,22,35,40,68,116,126,134–136,141]. In reporting outcomes, overall survival (OS) was the sole reported outcome in 38 studies, recurrence or recurrence free survival (RFS) in 17 studies, progression free survival (PFS) in four studies and disease-free survival (DFS) in two studies. Ten studies used a combination of these prognostic indicators. Performance metrics were variable, with papers reporting sensitivity (11), specificity (11), accuracy (22), positive predictive values (PPV) (4) and negative predictive values (NPV) (4), area under the receiver operator curve (AUROC) (31), C-index (13), Matthew's correlation coefficient (MCC) (2) and F1-score (2). When comparing groups following predictions or unsupervised clustering, 14 studies (19.72 %) reported hazard ratios or *p*-values. Eighteen studies (25.35 %) were externally validated, while 48 (67.61 %) were internally validated only through cross-validation or segmentation of the dataset. Five studies (6.85 %) were not internally or externally validated, all of which utilized unsupervised or semi-supervised methods.

3.2. Cervical cancer

There were 41 studies reporting on cervical cancer. Twenty five studies (60.98 %) used only hospital datasets, 15 (36.59 %) used database data and one study (2.44 %) used both. The median sample size was 189 with an interquartile range of 189. Eighteen studies (43.90 %) used clinicopathological data, of which 14 (34.15 %) used only clinicopathological data, two (4.88 %) used clinicopathological and genomic data and two (4.88 %) used clinicopathological and radiomic data. Twelve studies (29.27 %) used genomic data only and 11 studies (26.83 %) used radiomic data only.

In terms of ML methods, 14 studies (34.15 %) used RF, eight (19.51

%) used SVM, six (14.63 %) used ANN, six (14.63 %) used regression analysis and one (2.44 %) used DT. Eleven studies compared ML methods without a clear pattern of superiority. Eight studies compared ML methods to non-ML prediction approaches [48,65,97,101,107,148,157], and all showed superiority. Eighteen papers used only OS as their endpoint, while 11 used RFS, two used DFS and one used PFS. Nine studies used a combination of OS and PFS or RFS. Performance metrics were variable, with papers reporting sensitivity (4), specificity (4), accuracy (6), precision (2), recall (2), PPV (1) and NPV (1), AUROC (19), C-index (7) and F1-score (2). Seven studies were externally validated in a different data set, 33 were internally validated and one was not validated.

3.3. Uterine cancer

There were 28 studies reporting on uterine cancer. Nine studies (32.14 %) used only hospital datasets, 15 (53.57 %) used database data and four studies (14.29 %) used both. The median sample size was 538 with an interquartile range of 691. Fifteen studies (53.57 %) used clinicopathological data, of which 11 (39.29 %) used only clinicopathological data, three (10.71 %) used clinicopathological and genomic data and one (3.57 %) used clinicopathological and radiomic data. Twelve studies (42.85 %) used genomic data only, and one study (3.57 %) used radiomic and genomic data.

In terms of ML methods, seven studies (25 %) used RF, four (14.29 %) used regression analysis, three (10.71 %) used DT, three (10.71 %) used SVM and three (10.71 %) used ANN. Four studies compared ML to non-ML approaches and all showed superiority [79,88,93,123], with six studies comparing ML methods without clear patterns of superiority. Seventeen papers used only OS as their endpoint, five used RFS, and one used PFS. Five studies used a combination of OS and PFS or RFS. Performance metrics were again variable, with papers reporting sensitivity (3), specificity (3), accuracy (5), precision (1), recall (1), F1 score (1), AUROC (14) and C-index (3). Eight studies compared cluster groups statistically with a *p*-value only. Five studies were externally validated in a different data set, 22 were internally validated and one was not validated.

3.4. Risk of Bias (ROB) and applicability analysis

The PROBAST tool was applied to each included paper to assess ROB and applicability (Fig. 3; Table 2). 139 (100 %) studies were assessed as high ROB. All studies were high ROB in the analysis domain, with the studies' power (4.1), performance measure reporting (4.7) and use of univariate analysis to select predictor variables (4.5) being the most common concerns for ROB. 37.41 % of studies had high ROB in the outcome domain, 16.55 % had high ROB in the predictor domain, and 15.11 % had high ROB in the participants domain.

102 studies (73.38 %) had high concern for applicability, 17 (12.23 %) had unclear concern for applicability and 20 (14.39 %) had low concern for applicability. 1.44 % of papers had high concern for applicability in the outcome domain, 71.22 % had high concern for applicability in the predictors domain and 9.35 % had high concern for applicability in the participants domain.

4. Discussion

This review aimed to systematically review and evaluate the methodologies used when applying ML to prognosticate common gynecological malignancies. The results show some promise in the field of ovarian, uterine and cervical cancers, and demonstrated discriminate predictive ability and superiority when individual studies compared these methods to non-ML methods. However, there were frequent methodological and reporting shortcomings that limit any conclusions that can be drawn in this review, which can be improved upon in future works.

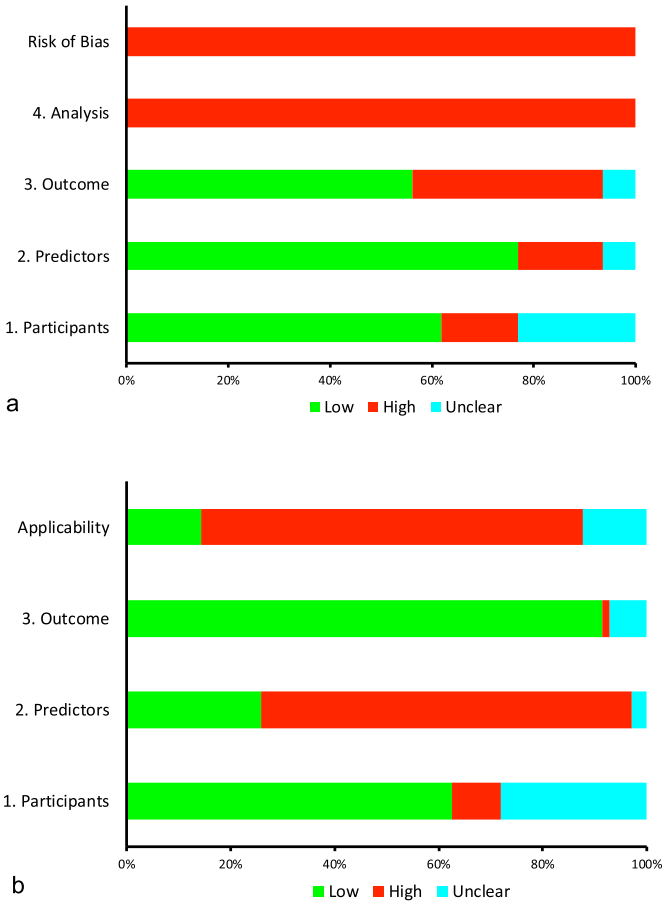

Fig. 3. Summarized Prediction model Risk Of Bias ASsessment Tool (PROBAST) results for all papers. 'Low' indicates low Risk Of Bias (ROB)/low concern regarding applicability. 'High' indicates high ROB/high concern regarding applicability. 'Unclear' indicates unclear ROB/unclear concern regarding applicability. a) Risk of Bias (ROB), including breakdown by four sub-domains – participants, predictors, outcomes and analysis. b) Applicability, including breakdown by three sub-domains – participants, predictors and outcomes.

4.1. Methodology analysis

4.1.1. Variable selection

The studies differed significantly in the number of variables tested for modelling. The studies that utilized clinicopathological predictors tended to use fewer numbers of variables that may often be collected as part of routine clinical activities. In most instances, these included patient age, FIGO stage, histological type, grade of tumor, biomarkers and residual tumor. Expectedly, the studies that tested thousands of predictor variables were often the ones that dealt with genomic data, including gene expression data, miRNA or mRNA expression profiles or a combination of those profiles. In addition, there was substantial variation in the reporting of how the variables were selected and/or retained in the final ML-based prediction model. It will likely be a long time before genomic markers can be utilized in a routine clinical environment, impacting their immediate translatability and applicability. However, the models utilizing clinicopathological and tumor biomarkers can be applied once validated externally, since these variables are more readily available.

4.1.2. ML methodology selection

There has been a distinct increase in the variety of ML methods used in the past decade. This may represent development of novel tools or increasing heterogeneity as the field expands. From the results, RF

Table 2
Prediction model Risk Of Bias ASsessment Tool (PROBAST) results for each paper. Results include assessments for both Risk of Bias (ROB) and Applicability as well as each sub-domain as listed. + indicates low ROB/low concern regarding applicability. - indicates high ROB/high concern regarding applicability. ? indicates unclear ROB/unclear concern regarding applicability.

Study	Risk of Bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Kappen et al. 1993 [20]	+	+	+	-	+	+	+	-	+
Kehoe et al. 2000 [21]	+	+	-	-	+	+	-	-	-
Snow et al. 2001 [22]	+	+	+	-	+	+	+	-	+
Ochi et al. 2002 [23]	+	+	+	-	+	+	+	-	+
Buciński et al. 2005 [24]	+	+	+	-	+	+	+	-	+
Oh et al. 2006 [25]	+	-	+	-	+	-	+	-	-
Tan et al. 2006 [26]	?	-	?	-	?	-	?	-	-
Buciński et al. 2007 [27]	+	?	+	-	+	+	+	-	+
Tan et al. 2008 [28]	?	+	+	-	?	-	?	-	?
Zhang et al. 2011 [29]	-	+	?	-	-	-	?	-	-
Ruan et al. 2012 [30]	-	?	-	-	-	-	?	-	-
Kim et al. 2013 [31]	+	+	?	-	+	-	+	-	-
Coveney et al. 2014 [32]	?	?	-	-	?	-	?	-	-
Kim et al. 2014 [33]	+	?	+	-	+	-	?	-	-
Tseng et al. 2014 [34]	+	+	+	-	+	+	+	-	?
Enshaei et al. 2015 [35]	+	+	+	-	+	+	+	-	+
Hassanzadeh et al. 2015 [36]	?	+	+	-	+	+	+	-	?
Liang et al. 2015 [37]	?	?	+	-	?	-	+	-	-
Zolbanin et al. 2015 [38]	+	+	+	-	+	+	+	-	+
Gligorijević et al. 2016 [39]	?	?	+	-	+	-	+	-	-
Ow et al. 2016 [40]	?	?	+	-	+	-	+	-	-
Qiu et al. 2016 [41]	-	+	+	-	+	-	+	-	-
Sun et al. 2016 [42]	?	+	+	-	+	-	+	-	-
Thomas et al. 2016 [43]	+	+	-	-	+	-	-	-	-
Verissimo et al. 2016 [44]	+	+	+	-	+	-	+	-	-
Wang et al. 2016 [45]	+	+	+	-	+	-	+	-	-
Liao et al. 2017 [46]	-	+	+	-	+	-	+	-	-
Matsuo et al. 2017 [47]	?	+	+	-	+	+	+	-	?
Obrzut et al. 2017 [48]	+	+	+	-	?	+	+	-	?
Cheng et al. 2018 [49]	+	+	-	-	+	-	+	-	-
Jing et al. 2018 [50]	+	+	-	-	+	-	?	-	?
Liao et al. 2018 [51]	+	+	+	-	+	+	+	-	+
Meng et al. 2018 [52]	+	+	-	-	+	-	?	-	-
Shinagare et al. 2018 [53]	+	+	+	-	+	+	+	-	+
Wang et al. 2018 [54]	?	-	-	-	+	-	+	-	-
Zhang et al. 2018 [55]	+	+	?	-	-	+	+	-	-
Zhou et al. 2018 [56]	+	-	-	-	+	-	+	-	-
Buttarelli et al. 2019 [57]	?	+	?	-	-	-	+	-	-
Dong et al. 2019 [58]	+	+	-	-	+	-	+	-	-
Hao et al. 2019 [59]	+	+	+	-	+	-	+	-	-
Kawakami et al. 2019 [60]	?	-	-	-	+	+	+	-	+
López-Reig et al. 2019 [61]	+	+	+	-	+	-	+	-	-
Lu et al. 2019(a) [62]	?	+	-	-	?	-	+	-	-
Lu et al. 2019(b) [63]	?	+	-	-	?	-	+	-	-
Mao et al. 2019 [64]	-	+	+	-	-	-	+	-	-
Matsuo et al. 2019 [65]	+	+	-	-	+	+	+	-	+
Mucaki et al. 2019 [66]	?	+	+	-	+	-	+	-	-
Octeau et al. 2019 [67]	+	+	+	-	+	-	+	-	-
Paik et al. 2019 [68]	?	+	-	-	+	+	+	-	+
Razak et al. 2019 [69]	?	+	+	-	+	-	+	-	-
Ruan et al. 2019 [70]	+	+	+	-	+	-	+	-	-
Shen et al. 2019 [71]	+	+	-	-	+	-	+	-	-
Upadhaya et al. 2019 [72]	?	-	+	-	+	-	+	-	-
Wang et al. 2019(a) [73]	?	+	-	-	+	-	+	-	-
Wang et al. 2019(b) [74]	+	+	+	-	+	-	+	-	-
Wang et al. 2019(c) [75]	+	+	+	-	+	-	+	-	-
Yin et al. 2019 [76]	+	+	-	-	+	-	+	-	-
Zhang et al. 2019 [77]	-	+	-	-	-	-	+	-	-
Cai et al. 2020 [78]	+	+	+	-	+	-	+	-	-
Casarin et al. 2020 [79]	-	+	-	-	-	+	+	-	-
Chen et al. 2020 [80]	+	+	+	-	+	-	+	-	-
Guerrero-Gimenez et al. 2020 [81]	?	+	+	-	+	-	+	-	-
Li et al. 2020 [82]	+	+	-	-	+	-	+	-	-
Liu et al. 2020 [83]	?	?	+	-	+	-	+	-	-
Mysona et al. 2020 [84]	-	?	+	-	-	+	+	-	-
Park et al. 2020 [85]	-	+	+	-	?	-	+	-	-
Praiss et al. 2020 [86]	+	+	+	-	?	+	+	-	?
Purohit et al. 2020 [87]	+	+	+	-	?	-	+	-	-
Qu et al. 2020 [88]	+	-	+	-	?	?	+	-	?
Reijnen et al. 2020 [89]	+	+	+	-	+	-	+	-	-
Takada et al. 2020 [90]	-	+	+	-	?	-	+	-	-

(*continued on next page*)

Table 2 (*continued*)

Study	Risk of Bias				Applicability			Overall	
	Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	ROB	Applicability
Veeraraghavan et al. 2020 [91]	-	-	-	-	?	-	+	-	-
Wallbillich et al. 2020 [92]	-	+	+	-	?	-	+	-	?
Wang et al. 2020 [93]	+	+	+	-	?	-	+	-	-
Xie et al. 2020 [94]	+	-	+	-	?	-	+	-	?
Zhang et al. 2020 [95]	?	-	-	-	?	-	+	-	-
Akazawa et al. 2021 [96]	+	-	+	-	+	+	+	-	+
Chai et al. 2021 [97]	+	+	+	-	-	-	+	-	-
Chang et al. 2021(a) [98]	+	-	?	-	?	+	+	-	?
Chang et al. 2021(b) [99]	+	+	-	-	?	-	+	-	-
Chen et al. 2021 [100]	+	-	+	-	?	-	+	-	-
Chu et al. 2021 [101]	+	-	+	-	?	-	+	-	-
Da-ano et al. 2021 [102]	+	-	+	-	?	-	+	-	-
Ding et al. 2021 [103]	?	-	+	-	?	+	+	-	?
Ferreira et al. 2021 [104]	+	-	+	-	+	-	+	-	-
Gargya et al. 2021 [105]	+	+	-	-	-	-	+	-	-
Grimley et al. 2021 [106]	-	+	+	-	-	+	+	-	-
Guo et al. 2021 [107]	+	+	-	-	-	+	+	-	-
Hoivik et al. 2021 [108]	+	+	+	-	+	-	+	-	-
Hsiao et al. 2021 [109]	+	+	+	-	+	-	+	-	-
Huo et al. 2021 [110]	?	+	?	-	?	-	+	-	-
Jajodia et al. 2021 [111]	-	+	+	-	+	-	+	-	-
Kaur et al. 2021 [112]	+	+	-	-	-	+	+	-	-
Kim et al. 2021(a) [113]	+	-	?	-	+	-	+	-	-
Kim et al. 2021(b) [114]	-	-	-	-	?	+	+	-	+
Kopanitsa et al. 2021 [115]	-	-	-	-	?	+	+	-	?
Laios et al. 2021 [116]	+	-	-	-	+	+	+	-	+
Li et al. 2021(a) [117]	+	+	+	-	?	-	+	-	-
Li et al. 2021(b) [118]	-	+	-	-	+	-	+	-	-
Ma et al. 2021(a) [119]	+	+	-	-	+	+	+	-	+
Ma et al. 2021(b) [120]	+	+	+	-	+	-	+	-	-
Mamrot et al. 2021 [121]	-	+	?	-	?	-	?	-	-
Mysona et al. 2021 [122]	+	+	+	-	?	-	+	-	-
Nakajo et al. 2021 [123]	+	+	-	-	+	-	+	-	-
Peng et al. 2021 [124]	+	+	-	-	+	-	+	-	-
Senthilkumar et al. 2021 [125]	?	+	+	-	?	-	+	-	-
Shannon et al. 2021 [126]	+	+	-	-	+	-	?	-	-
Sujamol et al. 2021 [127]	-	+	+	-	?	-	+	-	-
Tong et al. 2021 [128]	+	+	+	-	+	-	+	-	-
Wang et al. 2021 [129]	+	+	-	-	+	-	+	-	-
Yang et al. 2021 [130]	?	+	+	-	?	-	+	-	-
Zeng et al. 2021 [131]	?	+	-	-	+	-	+	-	-
Zhao et al. 2021 [132]	+	+	-	-	+	-	+	-	-
Zhou et al. 2021 [133]	?	+	+	-	?	?	+	-	?
Arezzo et al. 2022 [134]	?	+	+	-	?	+	+	-	+
Avesani et al. 2022 [135]	+	-	-	-	+	-	+	-	-
Belotti et al. 2022 [136]	+	+	-	-	+	-	+	-	-
Boehm et al. 2022 [137]	+	+	+	-	+	-	+	-	-
Carlini et al. 2022 [138]	?	-	+	-	?	-	+	-	-
Chen et al. 2022(a) [139]	+	+	-	-	?	?	+	-	?
Chen et al. 2022(b) [140]	+	+	+	-	+	-	+	-	-
Feng et al. 2022 [141]	-	+	-	-	?	+	+	-	?
Ladbury et al. 2022 [142]	+	+	+	-	?	+	+	-	+
Li et al. 2022 [143]	+	+	-	-	+	-	+	-	-
Liu et al. 2022(a) [144]	+	+	-	-	+	-	+	-	-
Liu et al. 2022(b) [145]	+	+	+	-	?	-	+	-	-
Liu et al. 2022(c) [146]	+	+	+	-	+	?	+	-	?
Meng et al. 2022 [147]	?	+	+	-	+	-	+	-	-
Nakajo et al. 2022 [148]	+	+	+	-	+	-	+	-	-
Piedimonte et al. 2022 [149]	+	+	+	-	+	+	+	-	+
Ruan et al. 2022 [150]	+	+	-	-	+	-	+	-	-
Tu et al. 2022 [151]	+	+	-	-	+	-	+	-	-
Vinklerová et al. 2022 [152]	-	+	-	-	+	-	+	-	-
Wang et al. 2022 [153]	+	+	-	-	+	-	+	-	-
Wei et al. 2022 [154]	+	+	-	-	+	-	+	-	-
Wu et al. 2022 [155]	+	+	-	-	+	-	+	-	-
Xu et al. 2022(a) [156]	?	+	-	-	+	-	+	-	-
Xu et al. 2022(b) [157]	+	+	+	-	+	+	+	-	+
Zhao et al. 2022 [158]	+	+	-	-	+	-	+	-	-

classifiers are the most commonly used in the field. These classifiers aim to generate ensembles of DTs, another commonly used method, from features that are randomly selected in the dataset [159]. This aims to improve upon DTs when handling highly complex data, which likely explains its common use in these studies. SVM and ANN classifiers were also commonly applied, which use nonlinear kernel functions like radial basis function, polynomial, sigmoid, tanh and ReLu. Hence, these classifiers can separate nonlinear features more accurately and yield high

classification performance. These ML methods involve pre-processing, feature extraction, statistical analysis and classification steps. The performance of such systems is subjective and time-consuming. DL methods do not involve these steps; the architecture extracts features and performs the classification. DL could be the way of the future without any of these limitations.

4.1.3. Endpoints

The endpoints selected for evaluating the predictive ability of these ML techniques were variable, in terms of both predicted outcome type (PFS, OS, DFS, RFS) and the method used. This variation included whether the methodology assessed continuous or binary survival, the time point selected (ranging from six months to five years), and if endpoints were selected before or after analysis. This variability raises questions as to what is clinically useful. It also prevents direct comparison of data, which may prevent translation of the best predictive algorithms into a clinical setting or meta-analysis of the field. Finally, a lack of consensus on endpoint selection and relevant methodology increases the risk of post-hoc selection, which was an observable risk in some studies analyzed using PROBAST. In the clinical prognostication literature, continuous OS and PFS are most often reported [160,161], which should likely be the aim for future ML prognostication models for clinical translatability. Thus, it would be useful for the establishment of consensus guidelines on endpoint selection and subtype, to prevent ROB, allow comparison, and facilitate clinical translation.

4.1.4. Statistical outcomes

The performance metrics also varied between studies. Whilst most studies reported discrimination adequately in the form of AUROC, C-index, or a sufficient combination of other measures (such as accuracy, precision, PPV, NPV, sensitivity and specificity), 27 % of studies insufficiently reported these measures. By only reporting single or limited discrimination measures such as accuracy, the predictive quality of the models outlined in these studies may not be readily assessed or as strong as the single measure would suggest [162]. Thirty one studies evaluated reported calibration measures, which greatly improves the reliability of these studies' conclusions [163]. With so few studies reporting calibration measures, most studies have ROB in their models, and are likely to be limited in real-world application and external validation. It would be ideal to compare study performance outcomes to identify which ML method, prediction variable set or data source is optimal. However, due to the heterogeneity in methodology, endpoints, data sources and statistical reporting, a quantitative meta-analysis or comparison of the outcomes reported could not meaningfully be performed in this setting. If future studies aim for homogenization toward the most clinically applicable parameters, a future review could aim to perform this meta-analysis.

4.1.5. ROB and applicability of reviewed papers

ROB and applicability were assessed using the PROBAST tool, and all 139 studies had high ROB. This identified the Analysis and Outcome domains as the two most common domains with high ROB. Within the Analysis domain, 89.92 % of studies not having sufficient events-per-variable (EPV) to have sufficient power in the model suggests that there is high risk of both bias and model failure when applied in the real-world setting. This phenomenon is particularly challenging in the ML space given the higher risk of overfitting; the PROBAST authors suggest even higher thresholds of EPVs should be applied to ML methods for this reason [19]. This poses challenges to study design and recruitment for the field, requiring either larger databases to be developed or longer recruitment periods within hospital settings. Using fewer features in model development would abate this risk; however this becomes particularly challenging in -omics study design. Furthermore, given that 34.53 % of studies did not have clear and appropriate inclusion or exclusion criteria, future works should aim to robustly define these cohorts. Relatedly, management of missing data points should be optimized through the use of specific datasets designed for the study or by imputation methods, to minimize the risk of missing data.

4.2. Strengths and limitations of review

This systematic review cast a wide net for available literature by utilizing multiple large databases and used a well-established ROB and applicability assessment tool in the PROBAST. However, we were unable to conduct meta-analysis of the data, given inconsistencies in endpoint selection, validation and performance metrics. Furthermore, although studies that compared ML and non-ML methods suggested superiority in ML methods, bias may exist given that the search strategy only included ML studies. As such, the question of superiority between ML and non-ML methods cannot be answered by this review. With respect to ROB analysis, the PROBAST tool has been developed through expert consensus and has been validated in a number of fields [19]; as such, it is the best tool available for this review. However, there are no ML-specific ROB analysis guidelines or tools that have been developed and validated, which is a limitation of this review. Consequently, the Enhancing the Quality and Transparency of Health Research Network (EQUATOR) is developing a PROBAST-AI tool to create a specific tool for this purpose [164]. Once validated, this tool or other ML-specific tools may be valuable for future reviews in this space.

4.3. Future direction

The pace at which ML-based models are being developed and proposed for routine clinical use is rapid. Future works need to develop more robust study design and reporting measures to facilitate this translation whilst minimizing ROB. The key areas of improvement are clear inclusion criteria, standardized endpoint selection, appropriate study power, management of overfitting risk, appropriate discrimination and calibration analysis, and transparent model reporting. Consensus guidelines may be useful for outlining methodological gold standards to pursue in future research in order to standardize metrics and therefore enable meta-analysis.

Our work has also identified a paucity of methods that utilized deep learning in this field. It could arguably be due to a lack of relevant data sources such as big data. When developed, these models are likely to be accurate and robust and do not require any manual extraction of features. Such systems can be installed on hospital servers or uploaded to the cloud, and accurate predictions can be obtained immediately based on information (clinicopathological, radiomic, proteomic and genomic) routinely collected by health service providers. However, such models that have widespread implications should be developed with input from data scientists, clinicians, ethicists, economists, implementation scientists and potential consumers. The models should be explainable to patients if we are to improve clinical outcomes.

5. Conclusions

It has been shown that the literature features ML models that may improve patient outcomes in the future with discriminate benefits over current methods; however, concerns regarding ROB and applicability exist with the currently available literature. Genomic and clinicopathological predictor variables, in combination with RF and SVM ML methods, have been the most commonly applied tools to date. However, there is significant heterogeneity in the field, and in recent times, unique ML methods have seen increasingly frequent use. Evaluation of the ROB and applicability with PROBAST identified several shortcomings that need to be addressed to progress CAPs to clinical translation, including improved analysis reporting, external validation of currently published methods, optimization of the best ML approaches, improved transparency, and standardization of methodology, endpoint selection and outcome reporting. This review identified that consensus guidelines outlining methodological gold standards for future research would be

valuable for the field and may facilitate greater meta-analytical insight and recommendations in the future. Until these shortcomings are overcome, this review is unable to conclude if ML would currently be of clinical value for prognostic purposes in gynecological oncology. However, given the promise shown by some studies in the field and the rate at which this field is expanding, it is possible that tangible clinical translatability and improvements to patient outcomes could be achieved in the future.

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.

Data availability

The data presented in this study are available in the article or supplementary material.

Acknowledgments

The authors would like to gratefully acknowledge the literature search contributions made by Ms. Jacky Cribb and Ms. Kaye Cumming, and analysis contributions from Ms. Daisy Eunji Cho, Ms. Yukei Oo and Ms. Yu Jin Cha.

CRediT authorship contribution statement

Conceptualization, JS., HR., RA., RG., XT., XZ., YL., and SKC.; methodology, JS., HR., RA., RG., XT., XZ., YL., and SKC.; formal analysis, JS., HR., HWL., and SKC.; data curation, JS., HR., and SKC.; writing—original draft preparation, JS., and HR.; writing—review and editing, JS., HR., RA., HWL., RG., XT., XZ., YL., TG., and SKC.; project administration, JS., HR., and SKC.; funding acquisition, SKC. All authors have read and agreed to the published version of the manuscript.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Some of the researchers who completed this work were employed under the Australian government's Rural Health Multidisciplinary Training Program and had full independence for this project.

Appendix A. Brief note about the classifiers commonly used

Classifiers used:

1. Logistic regression

Logistic regression can be used to categorize continuous variables by setting a breaking point [165]. It computes the probability score of an event, between 0 and 1. It can also be applied for datasets containing more than two variables such as satisfaction level (high, medium or low).

2. Support vector machine (SVM)

It is a supervised machine learning algorithm developed by Vapnik [166]. It consists of two main components: hyperplanes and support vectors. Hyperplanes are the decision boundaries that divide the data points into different classes. SVM ensures that the position of the decision boundary is kept at the maximum distance from the closest data points of each classed, which are also known as the support vectors. SVM is suitable for high-dimensional medical data classification.

3. Decision tree (DT)

It aims to classify items in the dataset based on the associated features, without compromising the interaction effect between each item [167]. DTs begins with a series of questions (nodes) whereby the top of the DT is the question with the lowest entropy (degree of randomness), followed by questions with higher entropy that are connected to the branch of the previous question. The leaves are located at the bottom of the DT and they are the classes which the items will be classified into. Hence, DTs produce a hierarchy of questions that are arranged accordingly to their entropy.

4. Random forest (RF)

It is developed to address the limitation of DTs when handling highly complex data; DTs lose generalization accuracy when applied for data with increasing complexity. RF generates ensembles of DTs from features that are randomly selected in the dataset [159].

5. Ensemble (EL) classifier

Ensemble (EL) classifier is a predictive model that is constructed from multiple machine learning models [167]. There are two types of EL approach: boosting and bagging. Adaptive Boosting (AdaBoost) is used to improve the performance of a weak learner iteratively [167]. Bagging combines the output of various learned classifiers into a single prediction. EL classifiers can be used as an alternative to improve other machine learning models like SVM and RF.

6. K-means clustering

K-means clustering was introduced to find an optimal partition in a dataset containing a large number of samples [167]. The letter 'K' refers to the number of clusters that we are interested in defining within a dataset. As a result, K number of centroids will be randomly placed in the dataset. The cluster indicates the collection of data points aggregated together due to certain similarities.

7. Artificial neural network (ANN)

Artificial neural network (ANN) is inspired by biological neurons and attempts to imitate how neurons transmit signals to one another [167]. Every layer in an ANN contains a certain number of neurons and there are three types of layers: input, hidden and output layer. The input layer receives the raw signals in the form of vectors, the hidden layer processes the input signals and updates the weight of the connection that exists between the neurons, and the processed signals reach the output layer where classification takes place.

8. Naïve Bayes (NB)

The Bayes theorem is a method to calculate the probability of one class of variables based on the attributes involved. Hence, a Bayesian network can be applied when the dataset is surrounded by uncertainties; Bayesian network attempts to construct a graphical model by linking random variables together, based on their causal relationship derived from Bayes theorem [159]. In the case of a Naïve Bayes (NB) classifier, the attributes are assumed to be independent of each other. Therefore, a NB algorithm will exclude the probability of attributes occurring, when computing the probability for each class variable.

9. Least absolute shrinkage and selection operator (LASSO)

Least absolute shrinkage and selection operator (LASSO) was developed to improve the performance of regression models. It attempts to retain salient features of the dataset by shrinking or setting some of

the coefficients to zero [168]. Hence, it is often employed in feature selection and it can also perform regularization of machine learning models simultaneously.

10. Gradient boosting machine (GBM)

It is developed to improve the performance of weak learners by making use of the loss function. GBM is an iterative process where each iteration produces a new classifier based on the negative gradient of the loss function in the previous iteration [169]. Hence, the resulting prediction model of GBM is an ensemble of classifiers with minimal prediction error.

Appendix B. Supplementary data

Supplementary Table 1. Search strategies used for systematic review, by database. Supplementary Table 2. Summarized paper-specific data extracted from the literature. Supplementary data to this article can be found online at https://doi.org/10.1016/j.artmed.2023.102536.

References

[1] National Cancer Institute. What is cancer?. 2021.
[2] Bray F, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68(6):394–424.
[3] Corrado G, et al. Optimizing treatment in recurrent epithelial ovarian cancer. Expert Rev Anticancer Ther 2017;17(12):1147–58.
[4] Kyrgiou M, et al. Survival benefits with diverse chemotherapy regimens for ovarian cancer: meta-analysis of multiple treatments. J Natl Cancer Inst 2006;98 (22):1655–63.
[5] Jayson GC, et al. Ovarian cancer. Lancet 2014;384(9951):1376–88.
[6] Bottoni P, Scatena R. The role of CA 125 as tumor marker: biochemical and clinical aspects. Adv Exp Med Biol 2015;867:229–44.
[7] Bray F, et al. Incidence trends of adenocarcinoma of the cervix in 13 European countries. Cancer Epidemiol Biomarkers Prev 2005;14(9):2191–9.
[8] Cohen PA, et al. Cervical cancer. Lancet 2019;393(10167):169–82.
[9] Morice P, et al. Endometrial cancer. Lancet 2016;387(10023):1094–108.
[10] Vickers AJ. Prediction models in cancer care. CA Cancer J Clin 2011;61(5): 315–26.
[11] Chu C, et al. Prognosticating for adult patients with advanced incurable cancer: a needed oncologist skill. Curr Treat Options Oncol 2020;21(1):5.
[12] Nagy M, Radakovich N, Nazha A. Machine learning in oncology: what should clinicians know? JCO ClinCancer Informa 2020;4:799–810.
[13] Richter AN, Khoshgoftaar TM. A review of statistical and machine learning methods for modeling cancer risk using structured clinical data. Artif Intell Med 2018;90:1–14.
[14] Shrestha P, et al. A systematic review on the use of artificial intelligence in gynecologic imaging – background, state of the art, and future directions. Gynecol Oncol 2022;166(3):596–605.
[15] Fiste O, et al. Machine learning applications in gynecological cancer: a critical review. Crit Rev Oncol Hematol 2022;179:103808.
[16] Akazawa M, Hashimoto K. Artificial intelligence in gynecologic cancers: current status and future challenges – a systematic review. Artif Intell Med 2021;120: 102164.
[17] Liberati A, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. PLoS Med 2009;6(7):e1000100.
[18] Wolff RF, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170(1):51–8.
[19] Moons KGM, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019;170 (1):W1–w33.
[20] Kappen HJ, Neijt JP. Neural network analysis to predict treatment outcome. Ann Oncol 1993;4(SUPPL. 4):S31–4.
[21] Kehoe S, et al. Artificial neural networks and survival prediction in ovarian carcinoma. Eur J Gynaecol Oncol 2000;21(6):583–4.
[22] Snow PB, Brandt JM, Williams RL. Neural network analysis of the prediction of cancer recurrence following debulking laparotomy and chemotherapy in stages III and IV ovarian cancer. Mol Urol 2001;5(4):171–4.
[23] Ochi T, et al. Survival prediction using artificial neural networks in patients with uterine cervical cancer treated by radiation therapy alone. Int J Clin Oncol 2002;7 (5):294–300.
[24] Buciński A, et al. Clinical data analysis with the use of artificial neural networks (ANN) and principal component analysis (PCA) of patients with endometrial carcinoma. RepPractOncolRadiother 2005;10(5):239–48.
[25] Oh JH, et al. Proteomic biomarker identification for diagnosis of early relapse in ovarian cancer. J Bioinform Comput Biol 2006;4(6):1159–79.
[26] Tan TZ, Ng GS, Quek C. A prognosis tool based on hemostasis and genetic complementary learning. In: 2006 International Conference on Biomedical and Pharmaceutical Engineering; 2006.
[27] Bucinski A, et al. Evaluation of selected prognostic factors in patients with ovarian cancer applying artificial neural network analysis. AdvClinExpMed 2007; 16(4):527–32.
[28] Tan QH, et al. Evolutionary algorithm for feature subset selection in predicting tumor outcomes using microarray data. In: Mandoiu I, Sunderraman R, Zelikovsky A, editors. Bioinformatics research and applications; 2008. p. 426–33.
[29] Zhang KX, Ouellette BFF. CAERUS: predicting CAncER outcomes using relationship between protein structural information, protein networks, gene expression data, and mutation data. PLoS Comput Biol 2011;7(3).
[30] Ruan J, et al. In: Network-based classification of recurrent endometrial cancers using high-throughput DNA methylation data; 2012. p. 418–25.
[31] Kim D, et al. ATHENA: identifying interactions between different levels of genomic data associated with cancer clinical outcomes using grammatical evolution neural network. BioData Min 2013;6(1).
[32] Coveney C, et al. Exploration of ovarian cancer microarray data focusing on gene expression patterns relevant to survival using artificial neural networks. In: Ortuno F, Rojas I, editors. Proceedings Iwbbio 2014: International Work-Conference on Bioinformatics and Biomedical Engineering. Vols. 1 and 2; 2014. p. 116–23.
[33] Kim S, Park T, Kon M. Cancer survival classification using integrated data sets and intermediate information. Artif Intell Med 2014;62(1):23–31.
[34] Tseng C-J, et al. Application of machine learning to predict the recurrence-proneness for cervical cancer. Neural ComputApplic 2014;24(6):1311–6.
[35] Enshaei A, Robson CN, Edmondson RJ. Artificial intelligence systems as prognostic and predictive tools in ovarian cancer. Ann Surg Oncol 2015;22(12): 3970–5.
[36] Hassanzadeh HR, Phan JH, Wang MD. A semi-supervised method for predicting cancer survival using incomplete clinical data. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); 2015.
[37] Liang M, et al. Integrative data analysis of multi-platform cancer data with a multimodal deep learning approach. IEEE/ACM Trans Comput Biol Bioinformatics 2015;12(4):928–37.
[38] Zolbanin HM, Delen D, Hassan Zadeh A. Predicting overall survivability in comorbidity of cancers: a data mining approach. DecisSupport Syst 2015;74: 150–61.
[39] Gligorijevic V, Malod-Dognin N, Przulj N. Patient-specific data fusion for cancer stratification and personalised treatment. Pac Symp Biocomput 2016;21:321–32.
[40] Ow GS, Kuznetsov VA. Big genomics and clinical data analytics strategies for precision cancer prognosis. Sci Rep 2016;6:36493.
[41] Qiu Y, et al. Early prediction of clinical benefit of treating ovarian cancer using quantitative CT image feature analysis. Acta Radiol 2016;57(9):1149–55.
[42] Sun CY, et al. A chemotherapy response classifier based on support vector machines for high-grade serous ovarian carcinoma. Oncotarget 2016;7(3): 3245–54.
[43] Thomas J, Sael L. Maximizing information through multiple kernel-based heterogeneous data integration and applications to ovarian cancer. In: Proceedings of the Sixth International Conference on Emerging Databases: Technologies, Applications, and Theory. Jeju, Republic of Korea: ACM; 2016. p. 95–8.
[44] Veríssimo A, et al. DegreeCox – a network-based regularization method for survival analysis. BMC Bioinforma 2016;17(16):449.
[45] Wang X, et al. A network-pathway based module identification for predicting the prognosis of ovarian cancer patients. JOvarian Res 2016;9(1). 73-73.
[46] Liao X, et al. In: Prognostic evaluation method of ovarian granulosa cell tumor based on semi-supervised collaborative intelligence model. 10; 2017. p. 96–103.
[47] Matsuo K, et al. A pilot study in using deep learning to predict limited life expectancy in women with recurrent cervical cancer. AmJObstetGynecol 2017; 217(6):703–5.
[48] Obrzut B, et al. Prediction of 5-year overall survival in cervical cancer patients treated with radical hysterectomy using computational intelligence methods. BMC Cancer 2017;17(1).
[49] Cheng L, et al. A random forest classifier predicts recurrence risk in patients with ovarian cancer. Mol Med Rep 2018;18(3):3289–97.
[50] Jing R, et al. Ensemble methods with voting protocols exhibit superior performance for predicting cancer clinical endpoints and providing more complete coverage of disease-related genes. IntJGenomics 2018;2018.
[51] Liao X, et al. Prognosis evaluation of ovarian granulosa cell tumor based on co-forest intelligence model. JEngSciTechnolRev 2018;11:135–42.
[52] Meng J, et al. Texture analysis as imaging biomarker for recurrence in advanced cervical cancer treated with CCRT. Sci Rep 2018;8(1):11399.
[53] Shinagare AB, et al. High-grade serous ovarian cancer: use of machine learning to predict abdominopelvic recurrence on CT on the basis of serial cancer antigen 125 levels. J Am Coll Radiol 2018;15(8):1133–8.
[54] Wang X, et al. Prediction of candidate RNA signatures for recurrent ovarian cancer prognosis by the construction of an integrated competing endogenous RNA network. Oncol Rep 2018;40(5):2659–73.
[55] Zhang F, et al. Predicting ovarian cancer recurrence by plasma metabolic profiles before and after surgery. Metabolomics 2018;14(5):65.
[56] Zhou J, et al. Establishment of a SVM classifier to predict recurrence of ovarian cancer. Mol Med Rep 2018;18(4):3589–98.

[57] Buttarelli M, et al. A combined ANXA2-NDRG1-STAT1 gene signature predicts response to chemoradiotherapy in cervical cancer. JExpClinCancer Res 2019;38 (1). 279-279.
[58] Dong J, Xu M. A 19-miRNA support vector machine classifier and a 6-miRNA risk score system designed for ovarian cancer patients. Oncol Rep 2019;41(6): 3233–43.
[59] Hao J, et al. Interpretable deep neural network for cancer survival analysis by integrating genomic and clinical data. BMC Med Genomics 2019;12(10):189.
[60] Kawakami E, et al. Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers. Clin Cancer Res 2019;25(10):3006.
[61] López-Reig R, et al. Prognostic classification of endometrial cancer using a molecular approach based on a twelve-gene NGS panel. Sci Rep 2019;9(1):18093.
[62] Lu H, et al. A mathematical-descriptor of tumor-mesoscopic-structure from computed-tomography images annotates prognostic- and molecular-phenotypes of epithelial ovarian cancer. Nat Commun 2019;10(1):764.
[63] Lu T-P, et al. Developing a prognostic gene panel of epithelial ovarian cancer patients by a machine learning model. Cancers 2019;11(2):270.
[64] Mao Y, et al. Identification of a 26-lncRNAs risk model for predicting overall survival of cervical squamous cell carcinoma based on integrated bioinformatics analysis. DNA Cell Biol 2019;38(4):322–32.
[65] Matsuo K, et al. Survival outcome prediction in cervical cancer: cox models vs deep-learning model. Am J Obstet Gynecol 2019;220(4):381.e1–381.e14.
[66] Mucaki EJ, et al. Predicting responses to platin chemotherapy agents with biochemically-inspired machine learning. Signal Transduct Target Ther 2019;4 (1):1.
[67] Octeau D, et al. Outcome-related differences in gene expression profiles of high-grade serous ovarian cancers following neoadjuvant chemotherapy. Mol Cancer Res 2019;17(12):2422.
[68] Paik ES, et al. Prediction of survival outcomes in patients with epithelial ovarian cancer using machine learning methods. JGynecolOncol 2019;30(4). e65-e65.
[69] Razak E, Yusof F, Ahmad Raus R. Cancer relapse prediction from microRNA expression data using machine learning. J Mech Continua Math Sci 2019;(Special Issue (1)):365–73.
[70] Ruan J, et al. A novel algorithm for network-based prediction of cancer recurrence. Genomics 2019;111(1):17–23.
[71] Shen W-C, et al. Prediction of local relapse and distant metastasis in patients with definitive chemoradiotherapy-treated cervical cancer by deep learning from [18F]-fluorodeoxyglucose positron emission tomography/computed tomography. Eur Radiol 2019;29(12):6741–9.
[72] Upadhaya T, et al. Comparison of radiomics models built through machine learning in a multicentric context with independent testing: identical data, similar algorithms, different methodologies. IEEE TransRadiatPlasma MedSci 2019;3(2): 192–200.
[73] Wang H, Li G. Extreme learning machine cox model for high-dimensional survival analysis. Stat Med 2019;38(12):2139–56.
[74] Wang Q, et al. Prognostic potential of alternative splicing markers in endometrial cancer. Mol Ther Nucleic Acids 2019;18:1039–48.
[75] Wang S, et al. Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer. Radiother Oncol 2019;132:171–7.
[76] Yin F, et al. Predicting prognosis of endometrioid endometrial adenocarcinoma on the basis of gene expression and clinical features using Random Forest. Oncol Lett 2019;18(2):1597–606.
[77] Zhang H, et al. Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study. Eur Radiol 2019;29 (7):3358–71.
[78] Cai S, et al. A 10-gene prognostic methylation signature for stage I-III cervical cancer. Arch Gynecol Obstet 2020;301(5):1275–87.
[79] Casarin J, et al. Presence of glandular cells at the preoperative cervical cytology and local recurrence in endometrial cancer. Int J Gynecol Pathol 2020;39(6): 522–8.
[80] Chen P, et al. Identification of prognostic immune-related genes in the tumor microenvironment of endometrial cancer. Aging 2020;12(4):3371–87. Albany NY.
[81] Guerrero-Gimenez ME, et al. Galgo: a bi-objective evolutionary meta-heuristic identifies robust transcriptomic classifiers associated with patient outcome across multiple cancer types. Bioinformatics 2020;36(20):5037–44.
[82] Li Y, et al. A prognostic nomogram integrating novel biomarkers identified by machine learning for cervical squamous cell carcinoma. J Transl Med 2020;18(1): 223.
[83] Liu M, Li H. Estimation of heterogeneous restricted mean survival time using random forest. Front Genet 2020;11:587378.
[84] Mysona DP, et al. Clinical calculator predictive of chemotherapy benefit in stage 1A uterine papillary serous cancers. Gynecol Oncol 2020;156(1):77–84.
[85] Park S, et al. Magnetic resonance imaging features of tumor and lymph node to predict clinical outcome in node-positive cervical cancer: a retrospective analysis. Radiat Oncol 2020;15(1):86.
[86] Praiss AM, et al. Using machine learning to create prognostic systems for endometrial cancer. Gynecol Oncol 2020;159(3):744–50.
[87] Purohit S, et al. Senescence-associated secretory phenotype determines survival and therapeutic response in cervical cancer. Cancers 2020;12(10). Basel.
[88] Qu W, et al. Development and validation of a personalized survival prediction model for uterine adenosarcoma: a population-based deep learning study. Front Oncol 2020;10:623818.
[89] Reijnen C, et al. Preoperative risk stratification in endometrial cancer (ENDORISK) by a Bayesian network model: a development and validation study. PLoS Med 2020;17(5):e1003111.
[90] Takada A, et al. A multi-scanner study of MRI radiomics in uterine cervical cancer: prediction of in-field tumor control after definitive radiotherapy based on a machine learning method including peritumoral regions. Jpn J Radiol 2020;38 (3):265–73.
[91] Veeraraghavan H, et al. Integrated multi-tumor radio-genomic marker of outcomes in patients with high serous ovarian carcinoma. Cancers 2020;12(11). Basel.
[92] Wallbillich JJ, et al. Identification of a transcriptomic signature with excellent survival prediction for squamous cell carcinoma of the cervix. Am J Cancer Res 2020;10(5):1534–47.
[93] Wang Q, et al. Novel miRNA markers for the diagnosis and prognosis of endometrial cancer. J Cell Mol Med 2020;24(8):4533–46.
[94] Xie L, et al. Prognostic assessment of cervical cancer patients by clinical staging and surgical-pathological factor: a support vector machine-based approach. Front Oncol 2020;10:1353.
[95] Zhang Y, et al. Integrated profiles analysis identified a coding-non-coding signature for predicting lymph node metastasis and prognosis in cervical cancer. Front Cell Dev Biol 2020;8:631491.
[96] Akazawa M, et al. The application of machine learning for predicting recurrence in patients with early-stage endometrial cancer: a pilot study. Obstet Gynecol Sci 2021;64(3):266–73.
[97] Chai H, et al. Integrating multi-omics data through deep learning for accurate cancer prognosis prediction. Comput Biol Med 2021;134:104481.
[98] Chang CY, et al. An artificial immune system with bootstrap sampling for the diagnosis of recurrent endometrial cancers. Open Med 2021;16(1):237–45. Wars.
[99] Chang H, et al. Construction of a macrophage infiltration regulatory network and related prognostic model of high-grade serous ovarian cancer. J Oncol 2021; 2021:1331031.
[100] Chen HZ, et al. A CT-based radiomics nomogram for predicting early recurrence in patients with high-grade serous ovarian cancer. Eur J Radiol 2021;145:110018.
[101] Chu R, et al. Risk stratification of early-stage cervical cancer with intermediate-risk factors: model development and validation based on machine learning algorithm. Oncologist 2021;26(12):e2217–26.
[102] Da-ano R, et al. A transfer learning approach to facilitate ComBat-based harmonization of multicentre radiomic features in new datasets. PLOS ONE 2021; 16(7):e0253653.
[103] Ding D, et al. Machine learning-based prediction of survival prognosis in cervical cancer. BMC Bioinforma 2021;22(1):331.
[104] Ferreira M, et al. [(18)F]FDG PET radiomics to predict disease-free survival in cervical cancer: a multi-scanner/center study with external validation. Eur J Nucl Med Mol Imaging 2021;48(11):3432–43.
[105] Gargya P, Bálint BL. Histological grade of endometrioid endometrial cancer and relapse risk can be predicted with machine learning from gene expression data. Cancers 2021;13(17). Basel.
[106] Grimley PM, et al. A prognostic system for epithelial ovarian carcinomas using machine learning. Acta Obstet Gynecol Scand 2021;100(8):1511–9.
[107] Guo C, et al. Novel artificial intelligence machine learning approaches to precisely predict survival and site-specific recurrence in cervical cancer: a multi-institutional study. Transl Oncol 2021;14(5):101032.
[108] Hoivik EA, et al. A radiogenomics application for prognostic profiling of endometrial cancer. Commun Biol 2021;4(1):1363.
[109] Hsiao YW, et al. A risk prediction model of gene signatures in ovarian cancer through bagging of GA-XGBoost models. J Adv Res 2021;30:113–22.
[110] Huo X, et al. Identification of a six-gene signature for predicting the overall survival of cervical cancer patients. Onco Targets Ther 2021;14:809–22.
[111] Jajodia A, et al. Combination of radiomics and machine learning with diffusion-weighted MR imaging for clinical outcome prognostication in cervical cancer. Tomography 2021;7(3):344–57.
[112] Kaur I, et al. An integrated approach for cancer survival prediction using data mining techniques. Comput Intell Neurosci 2021;2021:6342226.
[113] Kim BW, et al. Machine learning for recurrence prediction of gynecologic cancers using lynch syndrome-related screening markers. Cancers 2021;13(22). Basel.
[114] Kim SI, et al. Machine learning models to predict survival outcomes according to the surgical approach of primary radical hysterectomy in patients with early cervical cancer. Cancers 2021;13(15). Basel.
[115] Kopanitsa G, et al. Lifestyle cancer survival predictors: influence of vegetarian diet on the relapse of endometrial cancer. Stud Health Technol Inform 2021;285: 193–8.
[116] Laios A, et al. Feature selection is critical for 2-year prognosis in advanced stage high grade serous ovarian cancer by using machine learning. Cancer Control 2021;28:10732748211044678.
[117] Li H, et al. Radiomic score as a potential imaging biomarker for predicting survival in patients with cervical cancer. Front Oncol 2021;11:706043.
[118] Li HM, et al. Development of MRI-based radiomics model to predict the risk of recurrence in patients with advanced high-grade serous ovarian carcinoma. AJR Am J Roentgenol 2021;217(3):664–75.
[119] Ma J, et al. Artificial intelligence based on blood biomarkers including CTCs predicts outcomes in epithelial ovarian cancer: a prospective study. Onco Targets Ther 2021;14:3267–80.
[120] Ma JH, et al. An 8-gene DNA methylation signature predicts the recurrence risk of cervical cancer. J Int Med Res 2021;49(5):3000605211018443.
[121] Mamrot J, Hall NE, Lindley RA. Predicting clinical outcomes using cancer progression associated signatures. Oncotarget 2021;12(8):845–58.

[122] Mysona DP, et al. Tumor-intrinsic and -extrinsic (immune) gene signatures robustly predict overall survival and treatment response in high grade serous ovarian cancer patients. Am J Cancer Res 2021;11(1):181–99.
[123] Nakajo M, et al. Application of a machine learning approach for the analysis of clinical and radiomic features of pretreatment [(18)F]-FDG PET/CT to predict prognosis of patients with endometrial cancer. Mol Imaging Biol 2021;23(5):756–65.
[124] Peng J, et al. A ferroptosis-related lncRNA signature predicts prognosis in ovarian cancer patients. Transl Cancer Res 2021;10(11):4802–16.
[125] Senthilkumar G, et al. Incorporating artificial fish swarm in ensemble classification framework for recurrence prediction of cervical cancer. IEEE Access 2021;9:83876–86.
[126] Shannon NB, et al. A machine learning approach to identify predictive molecular markers for cisplatin chemosensitivity following surgical resection in ovarian cancer. Sci Rep 2021;11(1):16829.
[127] Sujamol S, Vimina ER, Krishnakumar U. Improving recurrence prediction accuracy of ovarian cancer using multi-phase feature selection methodology. ApplArtifIntell 2021;35(3):206–26.
[128] Tong L, Wu H, Wang MD. Integrating multi-omics data by learning modality invariant representations for improved prediction of overall survival of cancer. Methods 2021;189:74–85.
[129] Wang Y, et al. Construction of a microenvironment immune gene model for predicting the prognosis of endometrial cancer. BMC Cancer 2021;21(1):1203.
[130] Yang H, et al. Identification and analysis of an epigenetically regulated five-lncRNA signature associated with outcome and chemotherapy response in ovarian cancer. Front Cell Dev Biol 2021;9:644940.
[131] Zeng H, et al. Integration of histopathological images and multi-dimensional omics analyses predicts molecular features and prognosis in high-grade serous ovarian cancer. Gynecol Oncol 2021;163(1):171–80.
[132] Zhao D, et al. Pan-cancer survival classification with clinicopathological and targeted gene expression features. Cancer Informat 2021;20:11769351211035137.
[133] Zhou X, Li W, Wen Z. Data analysis for risk prediction of cervical cancer metastasis and recurrence based on DCNN-RF. JPhys: ConfSer 2021;1813(1):012033.
[134] Arezzo F, et al. A machine learning approach applied to gynecological ultrasound to predict progression-free survival in ovarian cancer patients. Arch Gynecol Obstet 2022;306(6):2143–54.
[135] Avesani G, et al. CT-based radiomics and deep learning for BRCA mutation and progression-free survival prediction in ovarian cancer using a multicentric dataset. Cancers 2022;14(11). Basel.
[136] Belotti Y, Lim EH, Lim CT. The role of the extracellular matrix and tumor-infiltrating immune cells in the prognostication of high-grade serous ovarian cancer. Cancers 2022;14(2). Basel.
[137] Boehm KM, et al. Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer. NatCancer 2022;3(6):723–33.
[138] Carlini G, et al. Prediction of overall survival in cervical cancer patients using PET/CT radiomic features. ApplSci 2022;12(12):5946.
[139] Chen C, et al. The pathological risk score: a new deep learning-based signature for predicting survival in cervical cancer. Cancer Med 2023;12(2):1051–63.
[140] Chen S, et al. A risk model of gene signatures for predicting platinum response and survival in ovarian cancer. JOvarian Res 2022;15(1):39.
[141] Feng Y, et al. Clinical analysis and artificial intelligence survival prediction of serous ovarian cancer based on preoperative circulating leukocytes. JOvarian Res 2022;15(1):64.
[142] Ladbury C, et al. Characterizing impact of positive lymph node number in endometrial cancer using machine-learning: a better prognostic indicator than FIGO staging? Gynecol Oncol 2022;164(1):39–45.
[143] Li Y, et al. Establish of an initial platinum-resistance predictor in high-grade serous ovarian cancer patients regardless of homologous recombination deficiency status. Front Oncol 2022;12:847085.
[144] Liu L, et al. Construction of a six-gene prognostic risk model related to hypoxia and angiogenesis for cervical cancer. Front Genet 2022;13:923263.
[145] Liu S, et al. Radiomics model of 18F-FDG PET/CT imaging for predicting disease-free survival of early-stage uterine cervical squamous cancer. Cancer Biomark 2022;33(2):249–59.
[146] Liu T, et al. EOCSA: predicting prognosis of epithelial ovarian cancer with whole slide histopathological images. Expert SystApplic 2022;206:117643.
[147] Meng X, et al. A novel attention-mechanism based cox survival model by exploiting pan-cancer empirical genomic information. Cells 2022;11(9).
[148] Nakajo M, et al. Machine learning based evaluation of clinical and pretreatment (18)F-FDG-PET/CT radiomic features to predict prognosis of cervical cancer patients. Abdom Radiol 2022;47(2):838–47. NY.
[149] Piedimonte S, et al. Predicting recurrence and recurrence-free survival in high-grade endometrial cancer using machine learning. J Surg Oncol 2022;126(6):1096–103.
[150] Ruan T, et al. Identification of a novel epithelial-mesenchymal transition-related gene signature for endometrial carcinoma prognosis. Genes 2022;13(2). Basel.
[151] Tu J, et al. Tumor DNA methylation profiles enable diagnosis, prognosis prediction, and screening for cervical cancer. Int J Gen Med 2022;15:5809–21.
[152] Vinklerová P, et al. External validation study of endometrial cancer preoperative risk stratification model (ENDORISK). Front Oncol 2022;12:939226.
[153] Wang X, Ye L, Li B. Development of a genomic instability-derived lncRNAs-based risk signature as a predictor of prognosis for endometrial cancer. J Cancer 2022;13(7):2213–25.
[154] Wei G, et al. MRI radiomics in overall survival prediction of local advanced cervical cancer patients tread by adjuvant chemotherapy following concurrent chemoradiotherapy or concurrent chemoradiotherapy alone. Magn Reson Imaging 2022;91:81–90.
[155] Wu H, et al. Construction and validation of a prognostic model based on 11 lymph node metastasis-related genes for overall survival in endometrial cancer. Cancer Med 2022;11(23):4641–55.
[156] Xu J, et al. RNA 5-methylcytosine regulators contribute to metabolism heterogeneity and predict prognosis in ovarian cancer. Front Cell Dev Biol 2022;10:807786.
[157] Xu Z, et al. Machine learning of dose-volume histogram parameters predicting overall survival in patients with cervical cancer treated with definitive radiotherapy. J Oncol 2022;2022:2643376.
[158] Zhao C, et al. Glycosylation-related genes predict the prognosis and immune fraction of ovarian cancer patients based on weighted gene coexpression network analysis (WGCNA) and machine learning. Oxid Med Cell Longev 2022;2022:3665617.
[159] Han J, Kamber M, Pei J. Data mining: concepts and techniques. 3rd ed. Morgan Kaufmann; 2011.
[160] El Bairi K, et al. Prediction of therapy response in ovarian cancer: where are we now? Crit Rev Clin Lab Sci 2017;54(4):233–66.
[161] Bookman MA. In: Can we predict who lives long with ovarian cancer?. 125(S24); 2019. p. 4578–81.
[162] Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 2020;21(1):6.
[163] Steyerberg EW, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21(1):128–38.
[164] Collins GS, et al. Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. BMJ Open 2021;11(7):e048008.
[165] Sperandei S. Understanding logistic regression analysis. BiochemMed 2014;24(1):12–8.
[166] Vapnik V. The nature of statistical learning theory. 2nd ed. New York: Springer-Verlag; 2000.
[167] Duda RO, Hart PE, Stork DG. Pattern classification. 2001.
[168] Tibshirani R. Regression shrinkage and selection via the lasso. J R Stat Soc B Methodol 1996;58(1):267–88.
[169] Natekin A, Knoll A. Gradient boosting machines, a tutorial. Front Neurorobot 2013;7(21).