communications medicine

A Nature Portfolio journal

Article

Article in Press

https://doi.org/10.1038/s43856-026-01905-y

Overview of models for predicting risk of developing ovarian cancer: a systematic review

Check for updates

Anastasia Ardasheva[1], Karyna Atha[1], Ikraan Mohamed[2], Hannah Nentwich[3], Bea Harris Forder[4], Kezia Gaitskell[5] & Christiana Kartsonaki[6]

Abstract

Background Ovarian cancer is the deadliest gynaecological malignancy, and most cases are diagnosed at an advanced stage. There is no population-wide screening; models for predicting risk of ovarian cancer could be useful in decreasing mortality. We aimed to provide the first comprehensive overview of models for predicting the risk of developing ovarian cancer and their performance, risk of bias, and applicability.

Methods We performed a systematic search through MEDLINE and Embase to identify articles in which models for predicting risk of developing ovarian cancer were developed and/or validated. Model characteristics and measures of predictive ability were extracted and summarised for each model. The Prediction model Risk of Bias Assessment Tool (PROBAST) was used to assess risk of bias and applicability.

Results We identified twenty-eight studies representing twenty-one ovarian cancer risk prediction models. Fifteen risk prediction models were developed or validated in the general population, while six were developed for high-risk populations, defined by family history, and one for patients with endometriosis. Twenty-three studies that reported discrimination and/or calibration had moderate-to-good performance. Risk prediction models with the best predictive ability reported an area under the receiver-operating characteristic curve (AUC) of 0.86 (95% confidence interval (CI) 0.84–0.87) for 2-year and an AUC of 0.77 (95% CI 0.76-0.78) for 10-year ovarian cancer risk in the general population.

Conclusions Several models for predicting short- and long-term risk of ovarian cancer have been published, of which a few have been validated in other populations. Some have limited predictive performance. Some models may be suitable for use in primary care or specialty clinics; however, assessment of their utility and validation in diverse cohorts is warranted.

Plain language Summary

We aimed to identify in the published literature all models that have been developed to make such predictions, and any studies that have been done to evaluate how well such models predict in another population.From the identified articles, we extracted information on the models developed or evaluated, as well as any information on their ability to accurately predict the risk of developing ovarian cancer in the future.We identified 28 articles and 21 models. Some were good at predicting the risk of ovarian cancer, but before they can be considered for use in practice, they would need to be evaluated in the population in which the prediction is to be made. If performing well, they may be useful for identifying individuals at a high risk of developing ovarian cancer who may benefit from early detection and prevention strategies or from enrolment into studies assessing the benefit of new such strategies.

Ovarian cancer is the deadliest gynaecological malignancy with over 207,000 deaths reported in 2022 worldwide[1]. The timing of ovarian cancer diagnosis significantly affects prognosis, as the 5-year survival rate changes from 92% for an early-stage ovarian cancer to 29% for a late-stage disease[2]. The high ovarian cancer mortality rate is largely attributed to delayed diagnosis, as over 70% of ovarian cancer cases are identified at advanced stages[3,4].

To assess whether general population screening reduces ovarian cancer mortality, large ovarian cancer screening trials were conducted in the USA and the UK. The US-based Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial and the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) did not observe a change in ovarian cancer mortality in their screening groups based on cancer antigen (CA125) testing and transvaginal ultrasounds[5–7]. Although UKCTOCS identified more early-stage ovarian cancer, it did not result in mortality reduction[7]. However, these studies

[1]Medical Sciences Division, University of Oxford, Oxford, UK. [2]University of Leicester, Leicester, UK. [3]University Hospitals Sussex NHS Foundation Trust, Worthing, UK. [4]Royal Surrey County Hospital, Guildford, UK. [5]Cancer Epidemiology Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK. [6]Clinical Trials Service Unit and Epidemiological Studies Unit (CTSU), Nuffield Department of Population Health, University of Oxford, Oxford, UK. e-mail: christiana.kartsonaki@ndph.ox.ac.uk

targeted the general population and used tests that were available at the time. One of the ways to improve a screening strategy is by utilising genomic data and other disease risk factors to identify the population most at risk of developing a disease[8]. Currently, women who are known to have a high familial or genetic risk may be offered surveillance or prevention strategies, but decision-making in this context is challenging, and there is no general consensus on how to identify such women[9]. Additionally, if such strategies existed, they would also benefit women at a lower risk of ovarian cancer. Learning about lower disease risk can have a positive psychological impact, such as reducing cancer worry, and promotes continued participation in a screening programme[10]. Thus, risk prediction models for ovarian cancer were developed to assess whether they could improve screening and risk stratification, and inform decisions on prophylactic interventions. A cancer risk prediction model assesses the probability of cancer occurring in the future, whereas a diagnostic model looks at the probability of disease being present in the patient at a given time. For the purposes of this study, we focus only on prediction models, as we are interested in answering the question of what prediction models exist, what variables they use, and what their prediction capabilities are. Some authors refer to such models as 'prognostic', but have not used this term to avoid confusion with models aiming to predict outcomes after cancer diagnosis. We decided to exclude diagnostic models, as they answer the question of how well a cancer lesion can be identified in a patient, and do not have any effect on disease prevention.

Although several ovarian cancer risk prediction models have existed for a number of years, no overview of the different models exists. This review, being the first work to summarise existing ovarian cancer risk prediction models as far as we are aware, will help systemise them and provide an overview of their performance, applicability and utility.

Methods

A systematic literature search was conducted through MEDLINE (Ovid) and Embase (Ovid), with a temporal limit from the inception of each database to 29th January 2022. The search was re-run on 20th January 2023 to capture studies published in 2022 and 2023, and duplicates were removed. Databases were searched for articles on gynaecological cancers (ovarian, uterine/endometrial, cervical, vaginal and vulvar), their risk prediction models,and validation studies. The search strategy is available in Supplementary Data 1. The search was limited to English-language articles. Conference papers and abstracts were excluded. References of original articles, reviews and meta-analyses were searched prior to the final analysis. The review protocol is published on PROSPERO (CRD42022303085).

Study selection was blinded and conducted using the Rayyan software. Titles and abstracts were screened independently by two co-authors (A.A. and C.K.), and disagreements were resolved by discussion after full-text screening. The authors decided to split studies by cancer type and only report ovarian cancer studies in this review, with other cancers reported as separate studies. Data including study design, population type, model variables and outcomes were extracted by two co-authors (A.A. and I.M.). We grouped studies by the model that they developed, extended, or validated. We grouped population types into general population, high-risk population (populations with a substantially higher risk than the general population based on hereditary factors) and 'special' populations (that is, populations defined by the presence of a feature that is not a typical high-risk feature). We also grouped models by the time frame of the prediction, into short- and long-term risk.

Studies were assessed in accordance with the Prediction Model Risk of Bias Assessment Tool (PROBAST)[11]. Each study was evaluated in four domains - participants, predictors, outcome and analysis—each ranked as 'high', 'low' or 'unclear' for the risk of bias and 'low', 'high' or 'unclear' concern for applicability (Table 1). We have assessed development and validation studies according to the PROBAST guidelines specific to each study type. Rating was performed twice independently (A.A., K.A., H.N., B.F.) and reviewed by another author (C.K.) and disagreements were resolved by discussion.

Studies and their characteristics were grouped by the model they developed, extended, or validated.

Ethical approval was not required for this study because it is a systematic review based on previously published studies and does not involve human participants or unpublished human data.

Results

Study characteristics

The initial search of all gynaecological cancer risk prediction models, including ovarian, endometrial/uterine, vulvar and cervical cancers, identified a total of 5033 abstracts and titles. Thirty-seven ovarian cancer risk prediction studies were assessed for eligibility, and 28 studies describing 21 models[12–39] were included in the final review. The review process is documented with the PRISMA chart (Fig. 1). The main reported outcomes were similar across all studies, ranging from ovarian cancer diagnosis within a pre-defined time period to lifetime risk of ovarian cancer. Several models focused on shorter-term ovarian cancer prediction, stating an ovarian cancer diagnosis within one to two years from study enrolment as their primary outcome[17,20,26,34]. Studies were geographically dispersed: three studies were conducted in large global consortia—the Consortium of Investigators of Modifiers of BRCA1/2 (CIMBA)[19,25] and the Ovarian Cancer Association Consortium (OCAC)[14], three studies were conducted in the European Prospective Investigation into Cancer and Nutrition (EPIC) cohort comprised of ten European countries[21,24,30], one study included data from five European countries[16], twelve studies took place in the UK[12,17,18,23,26,28,29,32–34,36,37], seven studies in the USA[15,20,27,31,35,38,39], one in Poland[22], and one in China[13]. Studies' population types were distributed as follows: twenty-one[12,14–18,21–24,28–36,38,39] studies were conducted in the general population, six[19,20,25–27,37] in high-risk populations and one[13] in endometriosis patients. High-risk populations were defined based on family history or known pathogenic mutations in BRCA1/2. Whereas we intended to include any definition of high risk population, we found that all published models define high risk based on family history; thus, throughout this review high risk population will be defined as such. Table 2 summarises risk prediction models' descriptions, locations and population types.

Model development overview

Commonly used modelling methods included logistic regression (10 studies), Cox proportional hazards models (7 studies), complex segregation analysis (2 studies) and machine learning methods—a polygenic risk score with continuous shrinkage machine learning model based on Bayesian regression (1 study)[14] and a Light Gradient Boosting machine learning model (1 study)[13]. Some studies, such as those by Antoniou and Pfeiffer, combined estimates from regression with population age-specific estimates of ovarian cancer incidence, taking into account competing causes of death, to estimate absolute risks of developing ovarian cancer. After examining OC risk prediction models, we decided to group their variables into nine categories: (1) demographic factors, (2) anthropometric factors, (3) lifestyle factors, (4) family history, (5) reproductive history, (6) genetic data, (7) laboratory tests, (8) symptoms and (9) medical history/clinical features. Table 3 provides a summary of model variables. We extracted 27 studies in this format, as the Lentz et al.[20] study was a short-term (<1 year) prediction model; thus, it will be discussed separately. Common combinations of variables were: demographic, anthropometric, lifestyle, family history, reproductive history and symptoms (5 studies)[13,24,28,30,34], demographic, anthropometric, lifestyle, family history, reproductive history and genetic data (3 studies)[16,18,27], demographic, anthropometric, family history and genetic data (3 studies)[19,29,37], demographic, anthropometric, lifestyle, reproductive history and medical history/laboratory tests (2 studies)[17,22], demographic factors and laboratory tests (2 studies)[26,33] and genetic data alone (2 studies)[14,25]. All studies included risk factors from more than two risk categories, and most studies included demographic (25 studies, 16 models), anthropometric (14 studies, 10 models), lifestyle factors (6 studies, 4 models) and reproductive history (16 studies, 12 models). Discrimination was assessed using the area under the receiver-operating characteristic curve

https://doi.org/10.1038/s43856-026-01905-y

Table 1 | PROBAST rating

	Model name (if applicable)	Author, year	Risk of bias				Applicability			Overall	
			Participants	Predictors	Outcome	Analysis	Participants	Predictors	Outcome	Risk of bias	Applicability
	General population										
1	Harvard Cancer Risk	Kim et al. 2004	+	−	+	−	+	+	+	−	+
2		Rosner et al. 2005	+	−	+	−	−	+	+	−	−
3	QCancer	Hippisley-Cox et al. 2011	+	+	+	+	+	+	+	+	+
		Hippisley-Cox et al. 2013	+	+	+	-	+	+	+	+	+
		Collins et al. 2013	+	+	+	+	+	+	+	+	+
		Hippisley-Cox et al. 2015	+	+	+	+	+	+	+	+	+
4		Pfeiffer et al. 2013	+	+	+	+	−	+	+	+	+
5		Grewal et al. 2013	+	−	+	−	+	+	+	−	+
6		Urban et al. 2015	+	−	+	−	+	+	+	−	+
7	Risk Assessment for Ovarian Cancer Chart (RAFOCC)	Pieta et al. 2018	+	−	−	−	−	+	+	−	−
8		Funston et al. 2021	+	+	+	+	−	+	+	+	−
9	Polygenic risk score (PRS) 15	Yang et al. 2018	+	+	+	+	−	+	+	+	−
10	EOC: Epithelial-Tubo Ovarian Cancer Risk Model	Lee et al. 2021	+	−	+	+	−	+	+	+	−
		Lee et al. 2022	+	+	+	+	+	−	+	+	+
11	Bright Pink Assess Your Risk (AYR) Tool	Hibler et al. 2022	−	?	+	−	+	+	?	−	?
12	Li et al. model	Li et al. 2015	+	+	+	+	+	+	+	+	+
		Fortner et al. 2017	+	+	+	-	+	+	+	-	+
		Fortner et al. 2018	+	−	+	−	+	+	+	−	+
13	'Select and Shrink for Summary Statistics' (S4)	Dareng et al. 2022	+	+	+	+	+	+	+	+	+
14	Combination of Dareng et al. and Li et al.	Dite et al. 2023	+	+	+	+	+	+	+	+	+
15	BOADICEA: Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm	Jervis et al. 2015	+	−	+	+	+	+	+	−	+
	High-risk populations										
1	BOADICEA: Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm	Antoniou et al. 2008	+	−	+	−	+	+	+	−	+
2	PRS 17	Kuchenbaecker et al. 2017	+	−	+	+	+	+	+	−	+
3	PRS 30	Barnes et al. 2020	+	+	+	+	+	+	+	+	+
4		Clyde et al. 2016	+	+	+	+	+	+	+	+	+
5		Russell et al. 2017	−	−	+	−	+	+	+	−	+
6		Lentz et al. 2020	−	+	+	?	+	+	+	−	+
	Special population (endometriosis patients)										
1		Chao et al. 2022	+	+	+	−	+	+	+	−	+

+ indicates low risk of bias/concern regarding applicability; − indicates high risk of bias/concern regarding applicability; ? indicates unclear risk of bias/concern regarding applicability.

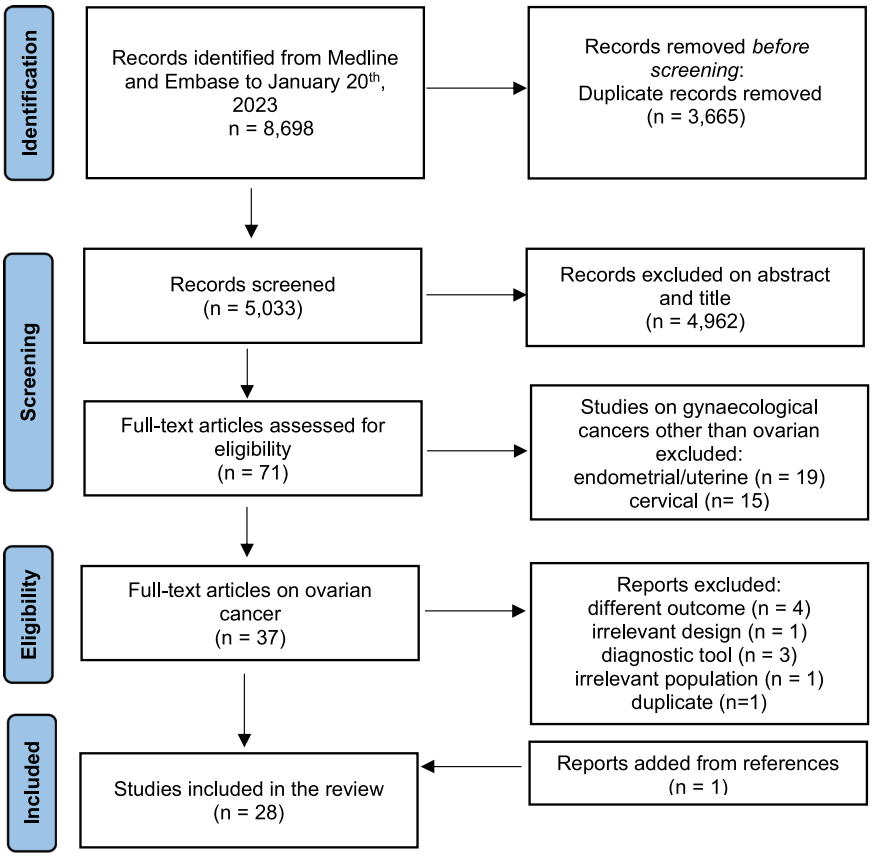


Fig. 1 | PRISMA flowchart of the systematic literature search conducted for gynaecological cancer risk prediction models.

(AUC) or C-statistic (and D-statistic for some), and model calibration was represented with a ratio of expected over observed cases or as a calibration slope (which measures the spread of predicted values). We used the terms as they were reported by the included studies. We do not make direct comparisons of numerical values of predictive performance as they are not generally comparable due to differences in the populations and outcomes.

Risk of bias

Risk of bias and applicability of studies describing models were assessed in accordance with the PROBAST tool[11] (Table 1). Out of twenty-eight included studies, 14 studies received a low risk of bias rating. Among the remaining studies, one was rated as high risk of bias in the 'participants' domain due to concerns about whether the study population was of appropriate design and inclusion and exclusion criteria were clearly defined. Eleven studies had high or unclear risk of bias in the 'predictors' domain, meaning that some variables were missing at the time of assessment or were assessed differently among participants. Ten studies had high or unclear risk-of-bias in the 'analysis' domain due to an inappropriate number of participants with the outcome, unclear categorisation of variables, or lack of information about model discrimination and/or calibration.

Evolution of ovarian cancer risk prediction models

Many of the ovarian cancer risk prediction models included in this review build upon one another or have been extended and/or updated over time. One such example is the BOADICEA model. Its first version[40,41] was designed for breast cancer and was later extended to assess the lifetime risk of both breast and ovarian cancer based on family history and BRCA1/2 mutation status[37,41]. Following the original version of BOADICEA, an updated version of the model has been published[19,23,25,29], designed specifically for ovarian cancer, that incorporated polygenic risk scores, BRCA1/2 status and family history, as well as a separate model for epithelial ovarian cancer (EOC model)[16,18]. This model has two versions and incorporates extended family history (beyond 1st degree relatives), reproductive history and genetic data for predicting the lifetime risk of ovarian cancer. BOADICEA and the EOC model are available as the online CanRisk Tool (www.canrisk.org).

Another defining risk prediction model, QCancer, was first derived for ovarian cancer[36], was then further enhanced and expanded to multiple cancers[34], and underwent external validation by an independent group[32]. Two versions of the QCancer model exist—the 2- and 10-year risk prediction, and both are available as an online tool (www.qcancer.org).[28,34]

Kuchenbaecker et al.[25] and Barnes et al.[19] developed polygenic risk scores in high-risk populations, and Yang et al.[23] and Dareng et al.[14] developed polygenic risk scores in the general population (Supplementary Table 2).

Another example of models building on existing models is the model developed by Li et al.[30] incorporating all categories of risk factors, which then served as a foundation for calculating absolute risk in the Fortner et al.[21,24] studies for assessing the utility of circulating biomarkers based on epidemiological risk factors. The Li et al. model was also combined with the polygenic risk score of Dareng et al.[14] by Dite et al.[12].

Predictive ability in the general population

As shown in Table 2, 21 studies (15 models) were conducted in the general population. Nine studies representing 6 ovarian cancer risk prediction models were ranked high on PROBAST, and 4 models were externally validated. Out of these models, the Hippisley-Cox et al.[34] QCancer model was developed in the primary care setting. The main outcome of QCancer was incident diagnosis of cancer within the first two years post-enrolment, as this time period was deemed representative of the length it takes for an existing cancer to manifest[34,42]. In addition to ovarian cancer, this model predicted the risk of other cancers, but we report discrimination and calibration for ovarian cancer only. Further external validation of QCancer[32] in the THIN database of UK primary care practices reported very good

https://doi.org/10.1038/s43856-026-01905-y

Table 2 | Summary of included ovarian cancer models

	Model name (if applicable)	Author, year	Model description	Prediction timeframe/ model output	Location/cohort
	General population				
1	Harvard Cancer Risk	Kim et al. 2004	Score based on demographic, lifestyle and medical history characteristics	10-year follow-up	USA
2		Rosner et al. 2005	Logistic regression based on demographic and reproductive history	Timeframe not specified	USA
3	QCancer	Hippisley-Cox et al. 2011	Symptom-based Cox proportional hazards model	2-year risk	UK
		Hippisley-Cox et al. 2013	Updated symptom-based multinomial logistic regression	2-year risk	UK
		Collins et al. 2013	An external validation of QCancer model	2-year risk	UK
		Hippisley-Cox et al. 2015	Cox proportional hazards model	10-year risk	UK
4		Pfeiffer et al. 2013	Absolute risk using population incidence rates and estimates from Cox proportional hazards model including reproductive and family history	Timeframe not specified	USA
5		Grewal et al. 2013	Logistic regression model for scoring OC symptoms in primary care		UK
6		Urban et al. 2015	Risk classifier based on Cox regression estimates		USA
7	Risk Assessment for Ovarian Cancer Chart (RAFOCC)	Pieta et al. 2018	Risk score based on odds ratios incorporating demographic, lifestyle factors and medical and social history	Timeframe not specified	Poland
8		Funston et al. 2021	Logistic regression based on demographic, lifestyle, reproductive history and CA125	1-year risk	UK
9	Polygenic risk score (PRS) 15	Yang et al. 2018	Logistic regression including polygenic risk score	Lifetime risk	UK
10	EOC: Epithelial-Tubo Ovarian Cancer Risk Model	Lee et al. 2021	Synthetic model incorporating genetic data, epidemiological risk factors and family history	5-year and lifetime risk	UK
		Lee et al. 2022	EOC model extension	Lifetime risk	UK, France, the Netherlands, Slovenia, Estonia
11	Bright Pink Assess Your Risk (AYR) Tool	Hibler et al. 2022	Algorithm based on family and personal medical history	Lifetime risk	USA
12	Li et al model	Li et al. 2015	Multivariable Weibull proportional hazards model	5-year risk	European Prospective Investigation into Cancer and Nutrition Cohort[a]
		Fortner et al. 2017	Conditional logistic regression including circulating OC early detection markers, medical history and lifestyle factors	-	
		Fortner et al. 2018	Conditional logistic regression based on levels of circulating anti-CA125 autoantibody, demographic and reproductive history	4-year risk	
13	'Select and Shrink for Summary Statistics' (S4)	Dareng et al. 2022	Machine Learning: Bayesian regression with continuous shrinkage polygenic risk score	Lifetime risk	Ovarian Cancer Association Consortium[a]
14	Combination of Dareng et al. and Li et al.	Dite et al. 2023	Cox regression including polygenic (36 SNPs) and clinical risk scores	10-year risk	UK
15	BOADICEA: Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm	Jervis et al. 2015	Segregation analysis of genetic signatures and family history of breast/ovarian cancer	Age-specific and lifetime risk	UK
	High-risk populations				
1	BOADICEA: Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm	Antoniou et al. 2008	Complex segregation analysis of family data	Age-specific and lifetime risk	UK
2	PRS 17	Kuchenbaecker et al. 2017	Cox regression including polygenic risk scores	Age-specific and lifetime risk	Consortium of Investigators of Modifiers of BRCA1/2[a]
3	PRS 30	Barnes et al. 2020	Cox regression incorporating polygenic risk score, demographic data and family history		

Table 2 (continued) | Summary of included ovarian cancer models

Model name (if applicable)	Author, year	Model description	Prediction timeframe/ model output	Location/cohort
4	Clyde et al. 2016	Logistic regression incorporating epidemiologic risk factors and single nucleotide polymorphisms	Timeframe not specified	USA
5	Russell et al. 2017	Logistic regression including blood-based biomarkers	2-year risk	UK
6	Lentz et al. 2020	Short-term prediction: Risk of Ovarian Cancer Algorithm combined with CA125 and HE4 serum levels		USA
Special population (endometriosis patients)				
1	Chao et al. 2022	Logistic regression and Light Gradient Boosting machine learning risk models	Timeframe not specified	China

[a]Details of the countries included in the consortia are provided in Supplementary Table 1.

discrimination with an AUC of 0.86 (95% CI 0.84–0.87) and a good calibration (represented as a graph). A longer-term Hippisley-Cox et al.[28] model predicting a 10-year risk of several cancers, including ovarian cancer, reported good discrimination for ovarian cancer prediction with an AUC of 0.769 (95% CI 0.76–0.778). The study published by Jervis et al.[29], describing the applicability of the BOADICEA model that incorporates polygenic risk scores and family history, was deemed to be low risk of bias in each domain of the PROBAST (Table 1), but did not report discrimination.

External validation of three other models developed and validated in the general population showed modest discrimination (Table 4). The Lee et al. 5-year OC risk prediction model had modest discrimination with an AUC of 0.61 (95% CI 0.58–0.64) and was well-calibrated with an expected-to-observed ratio of 1.05 (95% CI 0.94–1.16)[18]. The Pfeiffer et al. model predicting an absolute 5-year ovarian cancer risk on external validation reported an AUC of 0.59 (95% CI: 0.56–0.63) and an expected-over-observed ratio of 1.08 (95% CI: 0.97–1.19) in the US-based Nurses' Health Study cohort. It had a C-index of 0.55 (95% CI: 0.52–0.59) and slightly over-predicted risk, with an expected over-observed ratio of 1.35 (95% CI: 1.12–1.63) in the European EPIC cohort[30,35]. Dareng et al.[14] developed a risk prediction model based solely on genetic data and externally validated it in two different cohorts: the general population UK Biobank cohort and the high-risk population of the CIMBA consortium. They reported modest AUC values ranging between 0.538 and 0.629 for different ethnicities within each validation cohort, and calibration was not reported (Table 4).

High-risk populations

Six models were developed in high-risk populations, defined by family history. The BOADICEA model[37], that assesses risks of pathogenic BRCA1/2 and a lifetime breast and ovarian cancer risk based on family history on external validation had a good discrimination with a C-index of 0.77 (95% CI 0.73–0.82) and a good calibration with an expected to observed ratio of 0.87 (95% CI 0.70–1.08)[43] (Table 4). Two other studies ranked high by PROBAST included the polygenic risk score model, with a reported C-index of 0.604 (95% CI 0.582–0.626) for BRCA1 carriers and 0.667 (95% CI 0.636–0.699) for BRCA2[19,27].

The Lentz et al.[20] study conducted in the high-risk population examined the efficiency of early OC diagnosis in the Risk of Ovarian Cancer Algorithm, looking at serum levels of CA125 and HE4, followed by ultrasound every 4 months for women with abnormal results. This method was compared with the Standard of Care surveillance strategy, consisting of CA125 test and pelvic ultrasounds every 6 months. This study did not report discrimination or calibration. Additionally, the sample size was relatively small, including 149 patients in the test arm and 43 patients in the Standard of Care arm. Since the design of this risk prediction model was very different from other models, and based on the PROBAST, the overall risk of bias of this study is high in the participants and analysis domains, we did not extract it like other models. The authors claim that the derived model, which incorporates CA125 and HE4, had a lower average false positive rate than the Standard of Care arm[20].

Discussion

We have identified twenty-eight studies describing 21 ovarian cancer risk prediction models. Most models reported modest to high discrimination; however, several studies were ranked high for risk-of-bias due to concerns listed in PROBAST, although rigid interpretation of such criteria may not be desirable in all settings[11]. The overviewed models used different types of predictors that are not necessarily feasible or desirable to combine for the generation of one risk prediction model, as different models may not serve the same purpose and have different target populations. Different objectives include quantifying lifetime risk of ovarian cancer, which could inform screening and prevention strategies for individuals at high risk (e.g. salpingectomy, oophorectomy), long-term risk (over the next 5–20 years) to inform prevention or other decision-making, and using biomarkers and/or symptom combinations for short-term risk prediction or earlier diagnosis. Although we present the values of measures of discrimination and

https://doi.org/10.1038/s43856-026-01905-y

Table 3 | Summary of model variables

		General population													
		1	2	3				4	5	6	7	8	9	10	
		Kim et al.[39]	Rosner et al.[38]	QCancer				Pfeiffer et al.[35]	Grewal et al.[33]	Urban et al.[31]	Pieta et al.[22]	Funston et al.[17]	Polygenic risk score (PRS) 15	EOC model	
				Hippisley-Cox et al.[36]	Hippisley-Cox et al.[34]	Collins et al.[32]	Hippisley-Cox et al.[28]						Yang et al.[23]	Lee et al.[18]	Lee et al.[16]
Demographic factors	Age	X	X*	X	X	X	X	X[a]	X	X	X	X	X*	X*	X*
	Race									X		X		X	X
					postcode-related Townsend score, deprivation						residence (small vs big town)				
Anthropometric factors	BMI	X		X	X		X			X	X	X		X	X
	Height											X		X	X
Lifestyle factors	Smoking			X	X						X				
	Alcohol	X		X	X						X				
	Other										coffee drinking, sedentary lifestyle				
Family history	FH of OC/BC			X	X	X	X	X		X				X	X
	1st degree relative									X				X	X
	2nd degree relative									X				X	X
	Age/ ethnicity/ other data of relative with BC/OC									X				X	X
	Age/ ethnicity/ other data of unaffected relative													X	X
	Other cancers				All cancers										
Reproductive history	Age at menarche		X								X				

https://doi.org/10.1038/s43856-026-01905-y

Table 3 (continued) | Summary of model variables

		General population												
		1	2	3			4	5	6	7	8	9	10	
	Age at menopause		X							X				
	Parity	X	X				X		X	X			X	X
	Breastfeeding	X							X	X				
	HRT				X		X		X	X			X	X
	OCP	X	X		X	X	X		X	X			X	X
	Surgical procedures	Tubal ligation, hysterectomy	Tubal ligation, hysterectomy						Tubal ligation				Tubal ligation	Tubal ligation
	Other reproductive history				endometrial hyperplasia/polyp		Duration of menopausal hormone therapy use (MHT)			age at first pregnancy, age at first delivery, age at first mammogram, miscarriage, intervention during puerperium, time to weaning (for breastfeeding)	-		Endometriosis, breast tumour pathology (Oestrogen, progesterone, HER2 receptor, CK14, CK5/6 status)	Endometriosis, breast tumour pathology (Oestrogen, progesterone, HER2 receptor, CK14, CK5/6 status)
Genetic data	BRCA1/2 status								X				X	X
	Other genetic variants											15 SNPs	RAD51C, RAD51D and BRIP1	Added PALB2 to previous version
Laboratory tests									C-A125 and HE4		albumin, haemoglobin, platelets, CRP, CA125			
Symptoms				Deprivation, loss of appetite, weight loss, abdominal pain, abdominal distension, rectal bleeding,	Abdominal distention, abdominal pain, appetite loss, change in bowel habit, haematuria,	Abdominal distension, abdominal pain, loss of appetite, post-menopausal bleed,		Bloating, urinary frequency, rectal bleeding, postmenopausal bleeding, loss of appetite, abdominal pain,						

Table 3 (continued) | Summary of model variables

	General population											
	1	2	3			4	5	6	7	8	9	10
			postmeno-pausal bleeding, urinary frequency, diarrhoea, constipation, tiredness	indiges-tion, post-menopau-sal bleeding, weight loss	rectal bleed, weight loss		abdominal distension					
Medical history/ clinical features			Anaemia	Chronic pancreati-tis, chronic obstruc-tive pulmonary disease, type II diabetes, VTE, anaemia	Anaemia	Pre-vious diagno-sis of breast and/or cervical cancer			Depression	History of BC		

	General population							High-risk populations					Endometriosis patients
	11	12			13	14	15	1	2	3	4	5	1
	Hibler et al.[15]	**Li et al.[30] model**			**S4**	**Li et al.[30] and Dareng et al.[14]**	**BOADICEA**	**BOADICEA**	**PRS 17**	**PRS 30**	Clyde et al.[27]	Russell et al.[26]	Chao et al.[13]
		Li et al.[30]	Fortner et al.[24]	Fortner et al.[21]	Dareng et al.[14]	Dite et al.[12]	Jervis et al.[29]	Antoniou et al.[37]	Kuchenbaecker et al.[25]	Barnes et al.[19]			
Demographic factors		X[a]	X	X	X[b]	X[a]	X[a]	X[a]		X	X	X	X
Anthropometric factors		X	X			X					X		X
Lifestyle factors		X	X										
		X											
Family history	X	X	X				X	X		X	X		
	X	X					X	X		X	X		
	X												
	X						X						
	Pancreatic, endometrial, stomach, thyroid, kidney, sarcoma, adenocortical, brain, leukaemia												Endometrial cancer
Reproductive history		X	X			X					X		X
		X	X	X		X					X		

https://doi.org/10.1038/s43856-026-01905-y

Table 3 (continued) | Summary of model variables

	General population							High-risk populations					Endometriosis patients
	11	12			13	14	15	1	2	3	4	5	1
		X	X	X		X					X		
											X		
		X		X		X					X		
		X	X	X		X					X		
		Unilateral ovariectomy, hysterectomy	Oophrectomy, hysterectomy								Tubal ligation, hysterectomy		Benign cystectomy
	Abnormal breast biopsy, PCOS	menopausal status, number of full-term pregnancies (FTPs), age at first FTP, duration of breast-feeding, number of miscarriages, IUD use									Endometriosis		Usage of GnRHa before surgery, Mirena, Leiomyoma or adenomyoma, peritoneal endometriosis
Genetic data	X						X	X	X	X			
	Lynch syndrome (MSH2, MLH1, or EPCAM), Peutz-Jeghers Syndrome (STK11); BRIP1, RAD51C, RAD51D, PMS2 and MSH6				27,240 SNPs	36 SNPs from Dareng et al.	17 SNPs		17 SNPs	30 SNPs	17 SNPs		
Laboratory tests			CA125, CA15.3, HE4 and CA72.4	CA125 and anti-CA125 ab								Protein Z, Fibronectin, CRP, CA125	CA125, CA199
Symptoms													Dysmenorrhoea
Medical history/ clinical features	Chest radiation	Diabetes mellitus									Aspirin use		Tumour diameter, duration of disease, endometrial US

BMI body mass index, *FH* family history, *BC* breast cancer, *HRT* hormone replacement therapy, *OCP* oral contraceptive pill, *GnRH* gonadotropin-releasing hormone, *PCOS* polycystic ovarian syndrome, *US* ultrasound, *SNP* single nucleotide polymorphism, *IUD* intrauterine device, *VTE* venous thromboembolism.

[a]age is used as the timescale in a time-to-event model.

[b]age-specific incidence rates are used.

https://doi.org/10.1038/s43856-026-01905-y

Table 4 | Ovarian cancer (OC) risk prediction model performance

Model	Author, year	Population size (*n*)	Population age (range and/or mean [SD])	Model performance		
				Discrimination	Calibration	Other metrics
General population						
	Kim et al. 2004			*External validation of Harvard Cancer Index*	*External validation of Harvard Cancer Index*	
				C-statistic 0.59 (95% CI 0.56–0.62)	Risk Index-relative risk: reported as 'well calibrated'	
	Rosner et al. 2005	2,297,596: 472 OC cases	20–79	C-statistic 0.596 (95% CI 0.571–0.621)		
				Internal validation		
				Mann-Whitney statistic 0.594		
QCancer	Hippisley-Cox et al. 2011	1,158,723: 976 OC cases	30–84	*Internal validation*	*Internal validation*	*Internal validation*
			51 [15.5]			
				AUC 0.84 (0.83–0.86)	Graph: reported as 'good'	R^2 57.6 (54.8–60.4)
				D Statistic 2.38 (2.24–2.51)		
	Hippisley-Cox et al. 2013	1,240,864: 1279 OC cases	25–89	*Internal validation*	*Internal validation*	
			50.3 [17.5]			
				AUC 0.84 (95% CI 0.82–0.86)	Graph: reported as 'good'	
	Collins et al. 2013			*External validation of QCancer*	*External validation of QCancer*	*External validation of QCancer*
				AUC 0.86 (0.84–0.87)	Graph	R^2 59.9 (57.7–62.0)
				D Statistic 2.50 (2.38–2.62)		
	Hippisley-Cox et al. 2015	4,943,765: 6635 OC cases	25–84	*Internal validation*	*Internal validation*	*Internal validation*
				AUC 0.769 (95% CI 0.76–0.778)	Graph	R^2 29.1 (95% CI 26.8–31.4)
				D statistic 1.311 (95% CI 1.237–1.385)		
	Pfeiffer et al. 2013	199,973: 844 OC cases	55–74 (PLCO)	*External validation*	*External validation*	
			50–71 (NIH-AARP)			
				Nurses' Health Study	Nurses' Health Study	
				AUC 0.59 (95% CI 0.56–0.63)	E/O 1.08 (95% CI: 0.97–1.19)	
				EPIC	EPIC	
				C-index 0.55 (95% CI 0.52–0.59)	E/O 1.35 (95% CI 1.12–1.63)	
	Grewal et al. 2013	1272: 212 OC cases	40–80	AUC 0.883 (95% CI 0.853–0.912)		
	Urban et al. 2015	746: 210 OC cases	50–79			Hazard ratio 2.6
						Internal validation
						Hazard ratio 4.6
	Pieta et al. 2018	1346: 202 OC cases	18–80			72% sensitivity (95% CI 0.656–0.783), 75% specificity (95% CI 0.718–0.770)
	Funston et al. 2021	29,962: 279 OC cases	>18	AUC 0.938 (model 1)		
			55 [15]	AUC 0.941 (model 2)		
				Internal validation		
				AUC 0.935 (model 1)	*Internal validation*	
				AUC 0.935 (model 2)		

https://doi.org/10.1038/s43856-026-01905-y

Table 4 (continued) | Ovarian cancer (OC) risk prediction model performance

Model	Author, year	Population size (*n*)	Population age (range and/or mean [SD])	Model performance		
				Discrimination	Calibration	Other metrics
					Calibration slope 1.01 (95% CI 0.606–1.42) (model 1)	
					1.05 (95% CI 0.673–1.42) (model 2)	
Polygenic risk score (PRS) 15	Yang et al. 2018	2178: 750 OC cases	50–74	C-statistic 0.58 (95% CI 0.55–0.60)		
			63 [6.2]			
EOC model	Lee et al. 2021			*External validation*	*External validation*	
				AUC 0.61 (95% CI 0.58–0.64)	E/O = 1.05 (95% CI 0.94–1.16)	
	Lee et al. 2022		20–80			Risk <5%: 62% PALB2 carriers, risk 5–10%: 35% PALB2 carriers, risk >10%: 3% PALB2 carriers
	Hibler et al. 2022	143,657	≥18			Sensitivity 100%, specificity 89.9%
			29.66 [10.71]			
Li et al. model	Li et al. 2015	202,206: 324 OC cases	45–77.8, median 52.4	C-index 0.64 (95% CI 0.58–0.70)	E/O 0.91 (95% CI 0.82–1.01)	
				Internal validation	Calibration slope 0.88 (95% CI 0.64–1.11)	
				C-index 0.63 (95% CI 0.51−0.76)	*Internal validation*	
					E/O 0.90 (95% CI 0.74−1.08)	
	Fortner et al. 2017	2500: 590 OC cases	41–70	AUC 0.89 (95% CI 0.84–0.95)		
			56.3 [8.3]			
	Fortner et al. 2018	1072: 250 OC cases	30–78	AUC 0.84 (95% CI 0.76–0.92) (highest antibody tertile), 0.76 (95% CI 0.67–0.86) (lowest tertile)		
			58			
S4	Dareng et al. 2022	63,702: 23,564 OC cases, 40,138 controls		AUC 0.593		
				External validation		
				UK biobank		
				AUC 0.588 (European)		
				AUC 0.538 (East-Asian)		
				AUC 0.593 (African)		
				CIMBA		
				AUC 0.592 (BRCA1)		
				AUC 0.624 (BRCA2)		
Li et al. and Dareng et al.	Dite et al. 2023	189,171: 711 OC cases	40–69	Harrell's C-index 0.663 (95% CI 0.629–0.698)	Poisson regression 1.000 (95% CI 0.874–1.145)	
BOADICEA	Jervis et al. 2015	1548 OC cases	55.9 [9.4]		*Internal validation* E/O 0.95 (only mother) E/O 0.81 (only 1 sister) E/O 0.45 (mother and 1 sister)	
High-risk populations						
BOADICEA	Antoniou et al. 2008	2785 families: 130 OC cases	≥18	*External validation*	*External validation*	
				C-statistic 0.77 (95% CI 0.73–0.82)	E/O 0.87 (95% CI 0.70–1.08)	
PRS 17			>18	Harrell's C-index:		

https://doi.org/10.1038/s43856-026-01905-y

Table 4 (continued) | Ovarian cancer (OC) risk prediction model performance

Model	Author, year	Population size (*n*)	Population age (range and/or mean [SD])	Model performance		
				Discrimination	Calibration	Other metrics
	Kuchen-baecker et al. 2017	15,252: 2462 OC cases (BRCA1)				
				0.58 (95% CI 0.56 to 0.60) (BRCA1)		
		8211: 631 OC cases (BRCA2)				
				0.63 (95% CI 0.60 to 0.67) (BRCA2)		
PRS 30	Barnes et al. 2020	31,274: 2786 OC cases	≥18	Harrell's C-index: 0.604 (95% CI 0.582–0.626) (BRCA1)		*Internal validation*
			49 OC cases,			
			39 controls (BRCA1)	0.667 (95% CI 0.636–0.699) (BRCA 2)		Hazard ratio 1.28 (95% CI 1.06–1.55) (BRCA1)
			56.0 OC cases, 41.6 controls (BRCA2)			Hazard ratio 1.45 (95% CI 1.13–1.86) (BRCA2)
	Clyde et al. 2016	12,248: 4662 OC cases	57.58 [10.9] (cases)	*Internal validation*	*Internal validation*	
			56.2 [11.6] (controls)			
				AUC 0.664	Graph	
	Russell et al. 2017	80: 49 OC cases	50–74	AUC 0.949 (95% CI 0.917–0.982) (at <1year), 0.808 (95% CI 0.684–0.933) (1–2 years) (Model I)	27 out of 85 samples (98% specificity) (Model I)	*Internal validation*
			62.8 (58.7–67.3) (cases)			
			60.8 (58.4–65.8) (controls)	AUC 0.944 (95% CI 0.896-0.992) (at <1year), 0.638 (95% CI 0.522-0.755) (1-2 years) (Model II)	54 out of 158 samples (98% specificity) (Model II)	Root mean square error of logit scores:
						0.024 (Model I)
						0.020 (Model II)
Endometriosis patients						
	Chao et al. 2022	6809: 125 EAOC cases, 6684 non-EAOC	≥18	AUC 0.942 (95% CI 0.914–0.969) (machine learning)	C-index 0.891 (logistic regression)	
			36.4 ± 8.4			
				AUC 0.891 (95% CI 0.821–0.960) (logistic regression)		
				Internal validation		
				AUC 0.8758 (machine learning)		

SD standard deviation, *CI* confidence interval, *C-index/C-statistic* Concordance index, *AUC* area under receiver-operating curve, *EAOC* endometrial-associated ovarian cancer, *E/O* expected over observed ratio, R^2 estimated variation in time to ovarian cancer, *D* statistic discrimination.

calibration reported by the studies, we do not intend this to be a direct comparison of such values, as these cannot be directly compared due to the different target populations, prediction timeframes and study designs. For models that would be appropriate to use across different settings, external validation in the new setting would be desirable to ensure that the predictive ability of the model is maintained in the new setting.

Overall, fifteen studies describing nine models (nine in the general population, one model in the high-risk population) were deemed to have low risk-of-bias, and 19 had low concern for applicability and reported mostly modest discrimination. Despite an overall low risk-of-bias, some limitations of these studies were the lack of information about some model variables and concerns about model applicability. For example, the Pfeiffer et al.[35] model aimed to be applicable in the general population; however, this study was conducted in white women and may not be representative of the general US population, as 25% of the US population is non-white[44]. An external validation conducted by Li et al.[30] on the EPIC subcohort, comprised of Western and Central European countries, reported a modest C-index of 0.55 and a significant overprediction of ovarian cancer with an expected-over-observed ratio of 1.35 (95% CI 1.12–1.63)[30,45]. Future validation studies in diverse populations are warranted. The issue of population diversity was addressed by the Dareng et al. model, which relies solely on the genetic data and was validated in populations of different ethnicities. The discrimination against women of African ethnicity was higher than in European and East-Asian ethnicities, but the cohort was significantly smaller[14]. The disproportionate overrepresentation of white participants of European descent is a common problem in genome-wide association studies[46]. This is especially important because there are known differences in the epidemiology of ovarian cancer between ethnic groups. For example, Japanese Americans showed the strongest association of epithelial ovarian cancer with parity and oral contraceptive use, while for Latinas with postmenopausal hormonal use and age at menopause[47]. Additionally, there are differences in the distribution of histological types of ovarian cancer among various ethnic groups. For example, Asian Pacific Islanders have a significantly higher risk of ovarian clear cell carcinoma compared to the white population[48]. Thus, validating risk prediction models in diverse cohorts can help improve screening strategies in non-white women, though given that OC is a rare disease, obtaining sufficiently large cohorts of non-white women might be challenging.

QCancer models for the prediction of a 2-year and 10-year OC risk reported high AUC[28,32]. Both models were derived in a large primary care cohort in the UK and relied on electronic patient records (EPR). EPR is a robust way to collect patient data, but is largely affected by the healthcare setting, practice workload, variations between EPR interfaces, different coding systems, data extraction and processing tools and research dataset preparation and methodology[49]. Whether or not a patient presents to primary care also depends on their health literacy, socioeconomic status and accessibility to healthcare, which may affect symptom reporting[50]. Early-stage ovarian cancer symptoms are often vague, making it challenging for patients to judge when to seek medical attention[51]. The high AUC of the 2-year OC risk QCancer model on external validation makes it likely to be useful in the clinical setting, and variables included in this model, such as demographic and lifestyle factors, family and reproductive history and symptoms, are usually available through general practice records. This can help identify women at risk of ovarian cancer in general practice and guide further referral to specialty clinics; however, a validation in a non-UK cohort is needed if the model were to be used internationally[30,35], especially in countries with a different healthcare system

The BOADICEA model included age and family history data and reported a moderate discrimination on external validation[43]. It was the only model applied in both general and high-risk populations; though it lacked external validation in the general population[29]. As three-quarters of ovarian cancer cases are not carriers of pathogenic BRCA1/2 mutations[52], including data on detailed family history and other potentially pathogenic mutations is important in order to obtain more accurate predictions[53]. Inclusion of additional risk factors, such as anthropometric measures (BMI and height), a more detailed family history, reproductive history and additional pathogenic genes associated with ovarian cancer, led to a moderately high AUC and a good calibration[18]. This suggests that genetics in conjunction with readily available data such as demographics and other factors mentioned above can yield better prediction results and may be useful in high-risk populations. Some limitations of this model were the lack of information on family history in the UKCTOCS cohort, and that data on certain rarer pathogenic variants were limited. Additionally, the application of prediction models that use higher-cost tests like genetics is challenging in developing countries[54]. Approximately 20% of ovarian cancer cases are attributed to hereditary components, with 30 single-nucleotide polymorphisms and rare variants in 10 genes having been shown to be associated with ovarian cancer; however, there may be more undiscovered genetic factors. Additionally, models may serve different purposes: models with emphasis on potentially modifiable factors can be used to inform preventative strategies, while models based on genetics may provide information for lifetime risk and inform decisions for preventative strategies in high-risk individuals[55]. This reiterates that the applicability of a model depends on the setting, and all published models report good discrimination only in a very specific setting, making it challenging to recommend one universal model, and, as this review illustrates, at the moment, there is no one-size-fits-all approach to predicting ovarian cancer across varying populations.

Overall, only the QCancer model predicting a 2-year ovarian cancer risk reported low risk-of-bias and a very good discrimination (AUC > 0.8), although the 10-year-risk QCancer model reported an AUC just slightly below 0.8 and good calibration. Other models that reported high AUC (>0.8) values[13,21,24,26,33] had a high risk of bias according to PROBAST. One of these models was predicting endometriosis-associated ovarian cancer[13], so was not applicable in the general population. Two other models[21,24] were not validated, and another model[26] was based on a very small sample size. Discrimination was generally better in short-term risk prediction models. For example, the Russell et al. model, despite its small sample size, reported a significant decrease in AUC from 0.944 for ovarian cancer prediction within one year, to AUC of 0.638 between years one and two[26].

Ovarian cancer is a rare disease accounting for less than 5% of all cancer cases in women[56]. Attempts to perform population-level screening for ovarian cancer did not reduce mortality[5–7]; therefore, ovarian cancer risk prediction models developed on large, diverse patient cohorts could become a useful clinical tool. Even though screening the general population selected solely based on age has not improved outcomes, we speculate that identifying a well-selected target population in combination with further assessment might improve outcomes in the future. Whether the existing models are well-suited to predict ovarian cancer risk is a challenging question. It is not possible to directly compare measures of calibration, such as the C-statistic, between different models, and these can be used as merely indicators of each model's predictive ability. Discrimination and calibration are both important aspects of model performance that need to be evaluated[57]. Calibration is important for guiding clinical decisions, as poor calibration was found to reduce Net Benefit, a metric evaluating the relative consequences of false positives and negatives to assess the clinical utility of the model[58]. In certain situations where models are to be used for predicting risk, additional metrics may be helpful in determining the utility of a model in that setting, such as Net Benefit. However, neither calibration nor discrimination on their own are sufficient to make a clinical decision[59].

Some of the existing models have performed well in studies in their setting or target population and are available to use in practice, such as QCancer in general practice. One of the central arguments in the field of risk prediction is that prediction models should not be seen solely as a way to identify disease, but rather help with decision-making—be it referral for screening, prophylactic treatment, or surgery[60]. For example, one aim of BOADICEA was to provide refined estimates of absolute risk to women meeting more crude criteria of a high risk, to inform genetic counselling and decision-making[61]. Some evidence suggests that multimodal screening and prophylactic surgery are cost-effective in women at sufficiently high ovarian cancer risk[62]. Additionally, new clinical guidelines on managing women

with familial and genetic risk of ovarian cancer are currently in development[63]. Thus, in the future, it is important to explore how existing ovarian cancer risk prediction models could guide clinical decision-making.

Conclusion

Most existing ovarian cancer risk prediction models at present have moderate ability to predict long-term risk of ovarian cancer in both the general and high-risk populations, but some have better predictive ability, in particular those predicting short-term risk. Some models are more appropriate for use in general practice or specialty clinics, as some variables such as demographic, family and/or reproductive history are readily available; however, further evaluation and improvements of existing models in varied populations and settings may help with prevention or early detection of ovarian cancer in individuals with a higher risk of the disease.

Data availability

All extracted data are available in Supplementary Data 2.

Received: 10 July 2024; Accepted: 1 September 2026;
Published online: 05 November 2026

References

1. Bray, F. et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **74**, 229–263 https://doi.org/10.3322/caac.21834 (2024).
2. Reid, B. M., Permuth, J. B. & Sellers, T. A. Epidemiology of ovarian cancer: a review. *Cancer Biol. Med.* **14**, 9 (2017).
3. Brest A, et al. *SEER Cancer Statistics Review*, 1975–2017 (National Cancer Institute, 2017).
4. Lheureux, S., Gourley, C., Vergote, I. & Oza, A. M. Epithelial ovarian cancer. *Lancet* **393**, 1240–1253 https://doi.org/10.1016/S0140-6736(18)32552-2 (2019).
5. Buys, S. S. et al. Effect of screening on ovarian cancer mortality: the prostate, lung, colorectal and ovarian (PLCO) cancer screening randomized controlled trial. *JAMA* **305**, 2295–2303 https://doi.org/10.1001/jama.2011.766 (2011).
6. Pinsky, P. F. et al. Extended mortality results for ovarian cancer screening in the PLCO trial with median 15 years follow-up. *Gynecol. Oncol.* **143**, 270–275 (2016).
7. Menon, U. et al. Ovarian cancer population screening and mortality after long-term follow-up in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* **397**, 2182–2193 https://doi.org/10.1016/s0140-6736(21)00731-5 (2021).
8. Pashayan, N., Guo, Q. & Pharoah, P. D. Personalized screening for cancers: should we consider polygenic profiling? *Pers. Med.* **10**, 511–513 https://doi.org/10.2217/pme.13.46 (2013).
9. Redman, C., Duffy, S., Bromham, N. & Francis, K. Recognition and initial management of ovarian cancer: summary of NICE guidance. *Bmj* **342**, d2073 https://doi.org/10.1136/bmj.d2073 (2011).
10. Andrykowski, M. A. Psychological and behavioral impact of participation in ovarian cancer screening. *Diagnostics* **7**, 15 (2017).
11. Moons, K. G. M. et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann. Intern Med.* **170**, W1–w33 https://doi.org/10.7326/m18-1377 (2019).
12. Dite, G. S., Spaeth, E., Murphy, N. M. & Allman, R. A combined clinical and genetic model for predicting risk of ovarian cancer. *Eur. J. Cancer Prev.* **32**, 57–64 https://doi.org/10.1097/CEJ.0000000000000771 (2023).
13. Chao, X. et al. The application of risk models based on machine learning to predict endometriosis-associated ovarian cancer in patients with endometriosis. *Acta Obstet. Gynecol. Scand.* **101**, 1440–1449 https://doi.org/10.1111/aogs.14462 (2022).
14. Dareng, E. O. et al. Polygenic risk modeling for prediction of epithelial ovarian cancer risk. *Eur. J. Hum. Genet. EJHG* (2022).
15. Hibler, E. A. et al. Novel interactive tool for breast and ovarian cancer risk assessment (bright pink assess your risk): development and usability study. *J. Med. Internet Res.* **24**, e29124 https://doi.org/10.2196/29124 (2022).
16. Lee, A. et al. Enhancing the BOADICEA cancer risk prediction model to incorporate new data on RAD51C, RAD51D, BARD1 updates to tumour pathology and cancer incidence. *J. Med. Genet.* **59**, 1206–1218 https://doi.org/10.1136/jmedgenet-2022-108471 (2022).
17. Funston G, et al. Could ovarian cancer prediction models improve the triage of symptomatic women in primary care? A modelling study using routinely collected data. *Cancers* **13**, 2886 (2021).
18. Lee, A. et al. Comprehensive epithelial tubo-ovarian cancer risk prediction model incorporating genetic and epidemiological risk factors. *J. Med. Genet.* (2021).
19. Barnes, D. R. et al. Polygenic risk scores and breast and epithelial ovarian cancer risks for carriers of BRCA1 and BRCA2 pathogenic variants. *Genet. Med.* **22**, 1653–1666 (2020).
20. Lentz, S. E. et al. Development of a longitudinal two-biomarker algorithm for early detection of ovarian cancer in women with BRCA mutations. *Gynecol. Oncol.* **159**, 804–810 (2020).
21. Fortner, R. T. et al. Ovarian cancer early detection by circulating CA125 in the context of anti-CA125 autoantibody levels: results from the EPIC cohort. *Int. J. Cancer* **142**, 1355–1360 (2018).
22. Pieta, B. & Chmaj-Wierzchowska, K. Risk assessment for ovarian cancer. *Eur. J. Gynaecol. Oncol.* **39**, 517–523 (2018).
23. Yang, X. et al. Evaluation of polygenic risk scores for ovarian cancer risk prediction in a prospective cohort study. *J. Med. Genet.* **55**, 546–554 (2018).
24. Fortner, R. T. et al. Correlates of circulating ovarian cancer early detection markers and their contribution to discrimination of early detection models: results from the EPIC cohort. *J. Ovarian Res.* **10**, 20 (2017).
25. Kuchenbaecker, K. B. et al. Evaluation of polygenic risk scores for breast and ovarian cancer risk prediction in BRCA1 and BRCA2 mutation carriers. *J. Natl. Cancer Inst.* **109**, djw302 (2017).
26. Russell, M. R. et al. Novel risk models for early detection and screening of ovarian cancer. *Oncotarget* **8**, 785–797 (2017).
27. Clyde, M. A. et al. Risk prediction for epithelial ovarian cancer in 11 United States-based case-control studies: incorporation of epidemiologic risk factors and 17 confirmed genetic loci. *Am. J. Epidemiol.* **184**, 579–589 (2016).
28. Hippisley-Cox, J. & Coupland, C. Development and validation of risk prediction algorithms to estimate future risk of common cancers in men and women: prospective cohort study. *BMJ Open* **5**, e007825 https://doi.org/10.1136/bmjopen-2015-007825 (2015).
29. Jervis, S. et al. A risk prediction algorithm for ovarian cancer incorporating BRCA1, BRCA2, common alleles and other familial effects. *J. Med. Genet.* **52**, 465–475 (2015).
30. Li, K. et al. An epidemiologic risk prediction model for ovarian cancer in Europe: the EPIC study. *Br. J. Cancer* **112**, 1257–1265 (2015).
31. Urban, N. et al. Identifying post-menopausal women at elevated risk for epithelial ovarian cancer. *Gynecol. Oncol.* **139**, 253–260 (2015).
32. Collins, G. S. & Altman, D. G. Identifying women with undetected ovarian cancer: independent and external validation of QCancer(R) (Ovarian) prediction model. *Eur. J. Cancer Care* **22**, 423–429 (2013).
33. Grewal, K., Hamilton, W. & Sharp, D. Ovarian cancer prediction: development of a scoring system for primary care. *BJOG: Int. J. Obstet. Gynaecol.* **120**, 1016–1019 (2013).
34. Hippisley-Cox, J. & Coupland, C. Symptoms and risk factors to identify women with suspected cancer in primary care: derivation and validation of an algorithm. *Br. J. Gen. Pract.* **63**, e11–e21 (2013).
35. Pfeiffer, R. M. et al. Risk prediction for breast, endometrial, and ovarian cancer in white women aged 50 y or older: derivation and validation

from population-based cohort studies. *PLoS Med.* **10**, e1001492 (2013).
36. Hippisley-Cox, J. & Coupland, C. Identifying women with suspected ovarian cancer in primary care: derivation and validation of algorithm. *BMJ* **344**, d8009 (2011).
37. Antoniou, A. C. et al. The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions. *Br. J. Cancer* **98**, 1457–1466 https://doi.org/10.1038/sj.bjc.6604305 (2008).
38. Rosner, B. A., Colditz, G. A., Webb, P. M. & Hankinson, S. E. Mathematical models of ovarian cancer incidence. *Epidemiology* **16**, 508–515 (2005).
39. Kim, D. J., Rockhill, B. & Colditz, G. A. Validation of the Harvard Cancer Risk Index: a prediction tool for individual cancer risk. *J. Clin. Epidemiol.* **57**, 332–340 (2004).
40. Antoniou, A. C. et al. A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *Br. J. Cancer* **86**, 76–83 https://doi.org/10.1038/sj.bjc.6600008 (2002).
41. Antoniou, A. C., Pharoah, P. P., Smith, P. & Easton, D. F. The BOADICEA model of genetic susceptibility to breast and ovarian cancer. *Br. J. Cancer* **91**, 1580–1590 https://doi.org/10.1038/sj.bjc.6602175 (2004).
42. Jones, R., Latinovic, R., Charlton, J. & Gulliford, M. C. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using General Practice Research Database. *BMJ* **334**, 1040 https://doi.org/10.1136/bmj.39171.637106.AE (2007).
43. Ferris, J. S. et al. Abstract 4623: External validation of the BOADICEA model for predicting ovarian cancer risk: the Breast Cancer Family Registry. *Cancer Res.* **80**, 4623–4623 https://doi.org/10.1158/1538-7445.Am2020-4623 (2020).
44. Bureau USC. Quick Facts: U.S. Census Bureau; 2022. Available from: https://www.census.gov/quickfacts/fact/table/US/PST0452222023 (2023).
45. Ferlay, J. et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur. J. Cancer* **49** 1374–1403 https://doi.org/10.1016/j.ejca.2012.12.027 (2013).
46. Peterson, R. E. et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell* **179**, 589–603 https://doi.org/10.1016/j.cell.2019.08.051 (2019).
47. Sarink, D. et al. Racial/Ethnic differences in ovarian cancer risk: results from the multiethnic cohort study. *Cancer Epidemiol. Biomark. Prev.* **29**, 2019–2025 https://doi.org/10.1158/1055-9965.Epi-20-0569 (2020).
48. Korenaga, T.-R. et al. The elevated risk of ovarian clear cell carcinoma among Asian Pacific Islander women in the United States is not affected by birthplace. *Gynecol. Oncol.* **157**, 62–66 https://doi.org/10.1016/j.ygyno.2020.01.034 (2020).
49. Verheij, R. A., Curcin, V., Delaney, B. C. & McGilchrist, M. M. Possible sources of bias in primary care electronic health record data use and reuse. *J. Med Internet Res.* **20**, e185 https://doi.org/10.2196/jmir.9134 (2018).
50. Merriel, S. W. D. & Hamilton, W. Improving early cancer diagnosis in primary care. *Prescriber* **31**, 10–14 https://doi.org/10.1002/psb.1862 (2020).
51. Goff, B. A., Mandel, L. S., Melancon, C. H. & Muntz, H. G. Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. *JAMA* **291**, 2705–2712 https://doi.org/10.1001/jama.291.22.2705 (2004).
52. Jervis, S. et al. Ovarian cancer familial relative risks by tumour subtypes and by known ovarian cancer genetic susceptibility variants. *J. Med. Genet.* **51**, 108–113 https://doi.org/10.1136/jmedgenet-2013-102015 (2014).
53. Ramus, S. J. et al. Germline mutations in the BRIP1, BARD1, PALB2, and NBN genes in women with ovarian cancer. *J. Natl. Cancer Inst.* **107** https://doi.org/10.1093/jnci/djv214 (2015).
54. Kathryn, A. P. et al. Availability and funding of clinical genomic sequencing globally. *BMJ Glob. Health* **6**, e004415 https://doi.org/10.1136/bmjgh-2020-004415 (2021).
55. Kar, S. P. et al. Common genetic variation and susceptibility to ovarian cancer: current insights and future directions. *Cancer Epidemiol. Biomark. Prev.* **27**, 395–404 https://doi.org/10.1158/1055-9965.Epi-17-0315 (2018).
56. Momenimovahed, Z., Tiznobaik, A., Taheri, S. & Salehiniya, H. Ovarian cancer in the world: epidemiology and risk factors. *Int. J. Women's Health* **11**, 287–299 https://doi.org/10.2147/ijwh.S197604 (2019).
57. Van Calster, B. et al. Calibration: the Achilles heel of predictive analytics. *BMC Med.* **17**, 230 https://doi.org/10.1186/s12916-019-1466-7 (2019).
58. Van Calster, B. & Vickers, A. J. Calibration of risk prediction models: impact on decision-analytic performance. *Med. Decis. Mak.* **35**, 162–169 https://doi.org/10.1177/0272989x14547233 (2015).
59. Holmberg, L. & Vickers, A. Evaluation of prediction models for decision-making: beyond calibration and discrimination. *PLOS Med.* **10**, e1001491 https://doi.org/10.1371/journal.pmed.1001491 (2013).
60. Sachs, M. C., Sjölander, A. & Gabriel, E. E. Aim for clinical utility, not just predictive accuracy. *Epidemiology* **31**, 359–364 https://doi.org/10.1097/ede.0000000000001173 (2020).
61. Lee, A. et al. BOADICEA: a comprehensive breast cancer risk prediction modelincorporating genetic and nongenetic risk factors. *Genet. Med.* **21**, 1708–1718 https://doi.org/10.1038/s41436-018-0406-9 (2019).
62. Sroczynski, G. et al. A systematic review on cost-effectiveness studies evaluating ovarian cancer early detection and prevention strategies. *Cancer Prev. Res.* **13**, 429–442 https://doi.org/10.1158/1940-6207.Capr-19-0506 (2020).
63. NICE. Ovarian cancer: identifying and managing familial and genetic risk. In *Development* (National Institute for Health and Care Excellence, 2022).

Acknowledgements

We would like to thank Nia Roberts (Bodleian Library) for reviewing the search strategy.

Author contributions

Anastasia Ardasheva: Writing—Original Draft, Investigation, Methodology, Data Curation; Karyna Atha: Data Curation, Validation, Writing—Review & Editing; Ikraan Mohamed: Data Curation, Validation; Hannah Nentwich: Data Curation, Validation, Writing—Review & Editing; Bea Harris Forder: Data Curation, Validation, Writing—Review & Editing, Kezia Gaitskell: Writing—Review & Editing; Christiana Kartsonaki: Conceptualisation, Supervision, Investigation, Methodology, Writing—Review & Editing, Data Curation, Validation, Resources, Project administration.

Funding

C.K. and K.G. acknowledge funding from CRUK programme grant C16077/A29186.

Competing interests

C.K. was previously an Editorial Board Member for Communications Medicine but was not involved in the editorial review or peer review, nor in the decision to publish this article. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43856-026-01905-y.

Correspondence and requests for materials should be addressed to Christiana Kartsonaki.

Peer review information *Communications Medicine* thanks Lasai Barreñada and the other anonymous reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2026

This Article in Press is shared early to give you faster access to new research. It is citable and carries a permanent DOI. The final edited version will replace it automatically.