Received: 12 April 2024 | Revised: 3 June 2024 | Accepted: 27 June 2024

DOI: 10.1002/ijc.35092

RESEARCH ARTICLE

Innovative Tools and Methods

IJC INTERNATIONAL JOURNAL of CANCER | UICC

Role of artificial intelligence applied to ultrasound in gynecology oncology: A systematic review

Francesca Moro[1] | Marianna Ciancia[2,3] | Drieda Zace[4] | Marica Vagni[5] | Huong Elena Tran[6] | Maria Teresa Giudice[1] | Sofia Gambigliani Zoccoli[2,7] | Floriana Mascilini[1] | Francesca Ciccarone[1] | Luca Boldrini[6] | Francesco D'Antonio[8] | Giovanni Scambia[1,2] | Antonia Carla Testa[1,2]

[1]Dipartimento Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Fondazione Policlinico Universitario Agostino Gemelli, IRCCS, Rome, Italy

[2]Dipartimento Universitario Scienze della Vita e Sanità Pubblica, Università Cattolica del Sacro Cuore, Rome, Italy

[3]Dipartimento di Salute della Donna e del Bambino, Università degli studi di Padova, Padova, Italy

[4]Infectious Disease Clinic, Department of Systems Medicine, Tor Vergata University, Rome, Italy

[5]Istituto di Radiologia, Università Cattolica del Sacro Cuore, Rome, Italy

[6]Dipartimento di Diagnostica per Immagini, Radioterapia Oncologica ed Ematologia, Fondazione Policlinico Universitario Agostino Gemelli, IRCCS, Rome, Italy

[7]Department of Medical and Surgical Sciences for Mother, Child and Adult, University of Modena and Reggio Emilia, Azienda Ospedaliero Universitaria Policlinico, Modena, Italy

[8]Centre for Fetal Care and High-Risk Pegnancy, University of Chieti, Italy

Correspondence
Francesca Moro, Fondazione Policlinico Universitario Agostino Gemelli, IRCCS, Dipartimento Scienze della Salute della Donna, del Bambino e di Sanità Pubblica, Largo A. Gemelli 8, Rome, 00168, Italy.
Email: morofrancy@gmail.com; francesca.moro@policlinicogemelli.it

Abstract

The aim of this paper was to explore the role of artificial intelligence (AI) applied to ultrasound imaging in gynecology oncology. Web of Science, PubMed, and Scopus databases were searched. All studies were imported to RAYYAN QCRI software. The overall quality of the included studies was assessed using QUADAS-AI tool. Fifty studies were included, of these 37/50 (74.0%) on ovarian masses or ovarian cancer, 5/50 (10.0%) on endometrial cancer, 5/50 (10.0%) on cervical cancer, and 3/50 (6.0%) on other malignancies. Most studies were at high risk of bias for *subject selection* (i.e., sample size, source, or scanner model were not specified; data were not derived from open-source datasets; imaging preprocessing was not performed) and *index test* (AI models was not externally validated) and at low risk of bias for *reference standard* (i.e., the reference standard correctly classified the target condition) and *workflow* (i.e., the time between index test and reference standard was reasonable). Most studies presented machine learning models (33/50, 66.0%) for the diagnosis and histopathological correlation of ovarian masses, while others focused on automatic segmentation, reproducibility of radiomics features, improvement of image quality, prediction of therapy resistance, progression-free survival, and genetic mutation. The current evidence supports the role of AI as a complementary clinical and research tool in diagnosis, patient stratification, and prediction of histopathological correlation in

gynecological malignancies. For example, the high performance of AI models to discriminate between benign and malignant ovarian masses or to predict their specific histology can improve the diagnostic accuracy of imaging methods.

KEYWORDS
artificial intelligence, gynecology, machine learning, tumors, ultrasonography

What's New?

Artificial intelligence is increasingly used in advanced medicine and is now being applied to ultrasound imaging in gynecological oncology. However, a deeper understanding of the capabilities and limitations of AI would help improve the management of cancer patients from diagnosis to treatment. The current evidence analyzed in this systematic review of 50 studies supports the role of AI as a complementary research and clinical tool in diagnosis, patient stratification, and prediction of histopathological correlation in gynecological malignancies. AI applications are however still largely lacking for pathologies other than ovarian cancer.

1 | INTRODUCTION

Advances in gynecology ultrasound over the last two decades have led to significant increase in the detection of gynecological malignancies.[1–12] Currently, ultrasound represents the primary imaging modality for risk stratification of women presenting with ovarian masses.[13] Likewise, ultrasound assessment of women with endometrial or cervical cancer allows an accurate description of the topography of cancer invasion, allowing tailored surgical management based upon the imaging findings.[14,15]

The recent introduction of artificial intelligence (AI) in the field of diagnostic imaging has revolutionized the common approach to both diagnosis and prediction of prognosis in patients presenting with either benign or malignant conditions.[16] Indeed, AI has introduced new information that cannot be acquired from the standard clinical and radiological parameters.

In the last few years, there has been growing interest in the application of AI applied to imaging in gynecology. In particular, the use of AI has been extended to magnetic resonance imaging (MRI) and computed tomography with encouraging results.[17,18]

AI is created by feeding into predefined algorithms a multitude of relevant data, such as reasoning, learning, adaptation, sensory understanding, and interaction. This process is typically conducted by humans and requires the availability of extensive databases.[19]

Machine learning (ML) is a branch of AI that develops algorithms and statistical models to build computer systems that imitate human learning, without being explicitly programmed.[20] ML algorithms are trained on data to produce models and make decisions based on patterns observed.[21,22] The accuracy of the model increases as the input data increases. Recently, more evolved and combined neural networks have been used in deep learning (DL) to process complex data.

DL uses multilayer artificial neural networks that can remain a "black box" to the users and can automatically learn hierarchical representations of data, leading to the extraction of quantitative characteristics by digitally decoding images in order to identify even very small signs.[23]

In the field of gynecological imaging, AI models usually include clinical variables, imaging data, and radiomics features.

Radiomics is a technique used to extract, analyze, and interpret quantitative data from medical images.[24] The radiomics workflow involves different steps: image acquisition, tumor segmentation, quantitative features extraction from the tumor region, selection of the most informative features (i.e., statistical features indicative of intensity, textural features indicative of tissue architecture in terms of grey-level pixels), and analysis of their relationship with the outcome. The ultimate goal is the incorporation of quantitative imaging features into models in order to predict clinical endpoints (i.e., pathology diagnosis, staging, prognosis, treatment response).

Due to the growing application of AI in the past few years, some authors have explored the role of AI in gynecological oncology in systematic reviews and meta-analyses.[25–28] However, there is still a lack of synthesis of the available evidence regarding AI-based methods in ultrasound.

The aim of our review is to report the role of AI applied to ultrasound imaging in gynecology oncology.

2 | MATERIALS AND METHODS

2.1 | Search strategy

Web of Science, PubMed, and Scopus databases were searched to retrieve potential eligible articles, published until April 2, 2023. A search string for PubMed was structured consisting of Medical Subject Headings terms, keywords and free text words such as "radiomics" "ultrasound-based radiomics" "artificial intelligence" "machine learning" "deep learning" "Ultrasonography" "gynecology" "gynecological diseases" "endometrium" "uterus" "uterine" "ovary" "ovarian" "ovaries" "fallopian tube." The search was restricted to only humans and the English language. No other restrictions were used. The search string was adapted for use in the

other two electronic databases. The full search strategy for all databases can be found in Supplementary Note 1.

2.2 | Inclusion/exclusion criteria

Inclusion criteria were studies reporting the role of AI applied to ultrasound in gynecology oncology, specifically focusing on diagnosis of gynecological malignancies, image acquisition, quantification, segmentation, and location identification. Systematic reviews, nonempirical or animal studies, conference abstracts, editorials, commentaries, book reviews, and abstracts not accompanied by a full text were not considered eligible for inclusion in the present systematic review.

2.3 | Study selection

All studies retrieved from the search strategy were imported to RAYYAN QCRI software and duplicates removed. Two authors (F.M. and M.T.G) reviewed all abstracts independently. Agreement regarding potential relevance was reached by consensus. Full-text copies of the selected papers were obtained, and four reviewers (F.M., M.T.G., M.C., and S.G.Z.) independently extracted relevant data regarding study characteristics. We considered only papers reporting data on AI models applied to ultrasound imaging in the field of gynecological oncology. Inconsistencies were discussed by the reviewers and consensus reached or by discussion with the corresponding author. If more than one study was published for the same cohort with identical endpoints, the report containing the most comprehensive information on the population was included to avoid overlapping populations.

The reference lists of the included studies were hand-searched to look for additional studies.

When it was not possible to retrieve any full text online, we contacted the corresponding authors of the articles.

2.4 | Data extraction and analysis

Data extraction was performed by four researchers (F.M., M.C., M.V., and H.E.T.). A dedicated data extraction form was used to retrieve the following information for each eligible study:

(1) study identification: first author, title, publication year; (2) study characteristics: study period, country, design, disease, population; (3) the specific type of AI being assessed; (4) objective of the AI used and main findings (Table 1).

We performed a qualitative synthesis in the form of a narrative synthesis. The information retrieved from the included articles was categorized according to the type of AI assessed, and gynecological disease and was structured using Excel spreadsheets. The summary of findings was presented in a dedicated table including the specific AI used, the setting, the gynecological disease, and the objective for each of them (Table 1).

In studies including multiple developed models, the results of the best-performing model were reported in the "performance column" as the area under the receiver operating characteristic curve (AUC), otherwise as diagnostic accuracy, sensitivity and specificity, or positive and negative likelihood ratio and diagnostic odd ratio. For example, when an author developed more than one AI model, the model with the best performance was indicated. When the AUC was not reported in the article, the accuracy was considered and when both AUC and accuracy were not described, sensitivity and specificity were then indicated.

The indicated performance refers to the external or internal validation set; if no validation was performed, this information was not reported. For example, if an author developed a model validated in an external population, its performance refers to the results obtained from the external validation set; if the model was not externally validated, the performance obtained from the internal validation was reported. Where the model was neither externally nor internally validated, the performance refers to the data obtained from the developed model (Table 1).

2.5 | Quality assessment

The overall quality of selected studies was performed using the Quality Assessment Tool for Artificial Intelligence Centred Diagnostic Test Accuracy Studies (QUADAS-AI) criteria.[29] The specifics are listed in Supplementary. Table S1. The used criteria come from the extension and revision of QUADAS-2[30] and QUADAS-C[31] guidelines and comprises four domains (patient selection, index test, reference standard, flow, and timing) in the risk of bias. This new tool assesses each domain, providing a precise instrument to conduct reviews that evaluate AI-centered studies.

3 | RESULTS

3.1 | General characteristics

A total of 3118 articles were retrieved, 107 were assessed with respect to their eligibility for inclusion, and 50 studies were included in this systematic review (Figure 1). 37/50 (74.0%) studies were on ovarian masses or ovarian cancer, 5/50 (10.0%) on endometrial cancer, 5/50 (10.0%) on cervical cancer and 3/50 (6.0%) were on other gynecological malignancies.

The results of the quality assessment of the included studies using QUADAS-AI tool are presented in Supplementary Table S2. Most studies were at high risk of bias for subject selection (i.e., sample size or source were not specified; data was not derived from open-source datasets; imaging preprocessing was not performed and information on scanner model used to acquire images was not specified) and index test (i.e., the AI model was not tested in an external population in most articles) domains. However, there was generally a low risk of bias for reference standard (i.e., the reference standard reported in most studies correctly classified the target

TABLE 1 Data source for the 50 included studies.

First author	Year	Country	Assessment time	Sample size	Number of images	Gynecological pathologies	Objective of the study	Type of AI used	AI specifics	Model input	Performance	Type of validation
Van Calster B.	2007	Multicenter	NA	1066	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Relevance vector machine with a linear kernel	Clinical and ultrasound features	AUC = 0.95	Internal
Lucidarme O.	2010	Multicenter	NA	264	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Ovarian HistoScanningTM	Radiomics (statistical) features	Sensitivity = 0.98 Specificity = 0.88	External
Ahmadi E.	2011	Multicenter	NA	305	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Fuzzy decision tree	Clinical and ultrasound features	AUC = 0.91	Internal
Acharya U. R.	2012	NA	NA	20	2000	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Decision tree	Radiomics (textural, higher order spectra) features	Accucary = 0.95	Internal
Acharya U. R.	2012	NA	NA	20	2000	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Support vector machine	Radiomics (textural) features	Accuracy = 0.99	Internal
Faschingbauer F.	2013	NA	2005–2008	105	105	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Support vector machine	Radiomics (textural) features	Sensitivty = 0.69 Specificity = 0.86	Internal
Acharya U. R.	2013	NA	NA	20	2000	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Decision tree	Radiomics (textural) features	Accuracy = 0.97	Internal
Acharya U. R.	2014	NA	NA	20	2600	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL	Probabilistic neural network	Radiomics (textural) features	Accuracy = 0.99	Internal
Acharya U. R.	2014	NA	NA	20	2600	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL, ML	*DL: Probabilistic neural network (PNN); *ML: k-Nearest neighbor (KNN)	Radiomics (textural) features	*PNN accuracy = 1 *KNN accuracy = 1	Internal
Pathak H.	2015	NA	NA	NA	120	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Support vector machine	Radiomics (textural) features	Accuracy = 0.92	Internal
Aramendia-Vidaurreta V.	2016	Spain, Navarra	NA	145	145	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL	Multilayer perceptron networks	Clinical and radiomics (textural) features	AUC = 0.99	Internal
Acharya U. R.	2018	NA	NA	469	469	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Fuzzy forest (an ensamble fuzzy classifier)	Radon-transformed nonlinear features	AUC = 0.81	Internal
Wu C.	2018	China	2008–2018	NA	988	Ovarian cancer	HISTOLOGY (benign vs. malignant vs. borderline)	DL	GoogleNet(V3)	2D gray-scale images	AUC = 0.97	Internal
Martinez-Mas J.	2019	Belgium, Leuven	NA	NA	384	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Extreme learning machine with linear-sigmoid-Gaussian kernel	Geometric (fast Fourier Transform features	AUC = 0.88	Internal
Stukan M.	2019	Poland	2012–2017	290	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Multivariate regression (BIC criterion)	Ultrasound and clinical features	AUC = 0.98	Internal
Nero C.	2020	Italy, Rome	2013–2017	255	NA	Ovarian cancer	GENETICAL MUTATION	ML	Automated ML pipelines	Radiomics (textural, fractal) features	Accuracy = 0.64	Internal

(Continues)

TABLE 1 (Continued)

First author	Year	Country	Assessment time	Sample size	Number of images	Gynecological pathologies	Objective of the study	Type of AI used	AI specifics	Model input	Performance	Type of validation
							DETECTION (BRCA status)					
Akter L.	2021	United States	NA	NA (PLCO)	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Random forest	Clinical and ultrasound features	Accuracy = 0.81	Internal
Akter L.	2021	United States	NA	NA (PLCO)	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	XGBoost	Ultrasound features	AUC = 0.99	Internal
Al-karawi D.	2021	UK	2005–2013	232	242	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Support vector machine (SVM)	Radiomics (statistical, textural, fractal) features	Accuracy = 0.89	Internal
Chiappa V.	2021	Italy	2017–2021	274	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Support vector machine (SVM)	Radiomics features (TRACE4©)	Accuracy = 0.91	Internal
Chiappa V.	2021	Italy	2017–2019	241	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Ensemble of support vector machine (SVM) (TRACE4©)	Radiomics features (TRACE4©)	*solid OMs AUC = 0.87 *cystic OM AUC = 0.88 *mixed OMs AUC = 0.89	Internal
Christiansen F.	2021	Sweden	2010–2019	758	3077	Ovarian cancer	HISTOLOGY Task 1. Benign vs. malignant Task 2. Benign vs. inconclusive vs. malignant	DL	Ensamble model of VGG16, ResNet50 and MobileNet.	2D gray-scale images	*Benign versus malignant: AUC = 0.95; *Benign versus inconclusive versus malignant = AUC = 0.96	Internal
Hussein I. J.	2021	Multicenter	NA	NA	125	Ovarian cancer	METHODOLOGY (target automatic segmentation) HISTOLOGY (benign vs. malignant)	ML	*ML for segmentation: a novel Viola–Jones Model. *ML for prediction: Support vector machine (SVM)	Radiomics (textural) features	*SEGMENTATION Ovarian tumor dataset accuracy = 0.79 *CLASSIFICATION Ovarian tumor dataset accuracy = 0.95	Internal
Jin J.	2021	China	2002–2016	127	469	Ovarian cancer	METHODOLOGY (target automatic segmentation)	DL	Context encoder network	2D gray-scale images	DSC = 0.87	Internal
Qi L.	2021	China, Tianjin	2013–2016	265	NA	Ovarian cancer	HISTOLOGY Task 1. Benign vs. malignant vs. borderline Task 2. Borderline vs. malignant	ML	Multivariate logistic regression	Radiomics (shape, statistical, textural, wavelet) and clinical features	*Task1: AUC = 0.91 *Task2: AUC = 0.89	Internal
Ștefan P.A.	2021	Romania	2017–2019	120	123	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	k-nearest neighbor	Radiomics (textural) features	Accuracy = 0.85	No validation
Wang H.	2021	China	2013–2016	265	279	Ovarian cancer	HISTOLOGY (benign vs. malignant vs. borderline)	DL	ResNet34	2D gray-scale images	AUC = 0.96	Internal
Arezzo F.	2022	Italy	2018–2019	64	NA	Ovarian cancer	PROGRESSION FREE SURVIVAL	ML	Random forest	Clinical features	AUC = 0.92	Internal

TABLE 1 (Continued)

First author	Year	Country	Assessment time	Sample size	Number of images	Gynecological pathologies	Objective of the study	Type of AI used	AI specifics	Model input	Performance	Type of validation
Chen H.	2022	China	2019	422	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL	ResNet-18	2D gray-scale and color Doppler images	AUC = 0.93	Internal
Gao Y.	2022	Multicenter	2003–2019	107.624	592,275	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL	DenseNet-121	2D gray-scale images	*External validation dataset1: AUC = 0.87 *External validation dataset2: AUC = 0.83	External
Hsu S.T.	2022	China, Taiwan	NA	587	1896	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL	Ensemble Convolutional neural networks (ResNet-18, ResNet-50, and Xception)	2D gray-scale images	Accuracy = 0.92	Internal
Li J.	2022	Multicenter	2015–2022	2021	6965	Ovarian cancer	HISTOLOGY (benign vs. malignant vs. borderline) METHODOLOGY (target automatic segmentation)	DL	*mass segmentor (U-Net with LKResnet-18) *type classifier (LKResnet-18)	2D gray-scale and color Doppler images	*mass segmentator dice score: dataset1: DSC = 0.92 dataset2: DSC = 0.91 *type classifier macro-F1 score: dataset1: macro-F1 score = 0.75 dataset2: macro-F1 score = 0.68	External
Srilatha K.	2022	India	NA	NA	150	Ovarian cancer	HISTOLOGY (benign vs. malignant)	DL	Convolutional neural networks	Radiomics (shape, textural) features	Accuracy = 0.98	Internal
Tang Z.P.	2022	China	2017–2021	154	NA	Ovarian cancer	HISTOLOGY (histopathological types of epithelial ovarian cancer—type I vs. type II)	ML	Logistic regression	Clinical and radiomics (intelligenceFoundry software) features	AUC = 0.87	Internal
Yin Q.	2022	China, Guangzhou	2008–2019	137	NA	Ovarian cancer	HISTOLOGY (benign vs. malignant)	ML	Upgraded logistic regression	Radiomics (textural) features	Accuracy = 0.97	No validation
Zou Y.	2022	USA	2017–2018	35	600	Ovarian cancer	HISTOLOGY (benign vs. malignant) METHODOLOGY (reconstruct the optical absorption distribution from photoacustic tomography data)	ML, DL	*DL: ResNet-18 model to extract ovarian tissue morphology features; Ultrasound-enhanced Unet model (US-Unet) to reconstruct the optical absorption distribution from photoacustic tomography (PAT) data. *ML to distinguish benign from malignant cases: logistic regression	*DL inputs: ResNet-18—2D gray-scale images; US-Unet—US image features from the ResNet-18 and the initial pressure; *ML inputs: %sO2 and total hemoglobin concentration (HbT)	*ML: AUC = 0.94	Internal
Chen L.	2023	China	NA	NA	1469	Ovarian cancer	METHODOLOGY Task 1. Improve image quality	DL	*Removal of paintings on the images: mask-guided generative	2D gray-scale images	*Removal of paintings on the images Structural similarity (SSIM) = 0.92;	Internal

(Continues)

TABLE 1 (Continued)

First author	Year	Country	Assessment time	Sample size	Number of images	Gynecological pathologies	Objective of the study	Type of AI used	AI specifics	Model input	Performance	Type of validation
							Task 2. Target automatic segmentation		adversarial network (MGGAN) (based on fast Fourier convolutions FFCs) *Auto-segmentation: U-Net		*Auto-segmentation accuracy = 0.76	
Michail G.	2007	Greece	NA	65	NA	Endometrial cancer	HISTOLOGY (benign vs. malignant)	ML	Logistic regression	Radiomics (statistical, textural) features	AUC = 0.96	No validation
Angioli R.	2013	Italy	2010–2012	675	NA	Endometrial cancer	HISTOLOGY (benign vs. malignant)	ML	Logistic regression	Clinical and ultrasound features	AUC = 0.92	Internal
Moro F.	2022	Multicenter	2016–2019	498	NA	Endometrial cancer	HISTOLOGY (high risk vs. others)	ML	Logistic regression	Clinical and ultrasound features	AUC = 0.90	External
Xu J.	2022	China	NA	NA	NA	Endometrial cancer	PREDICTION OF MIOMETRIAL INVASION (deep vs. shallow)	ML	Random forest	Clinical and ultrasound features	Accuracy = 0.98	External
Ruan H.	2023	China	2013–2021	1837	NA	Endometrial cancer	HISTOLOGY (benign vs. malignant)	ML	Logistic regression	Clinical and ultrasound features	AUC = 0.91	Internal
Jin X.	2020	China	2014–2018	172	8–15 (for each patient)	Cervical cancer	HISTOLOGY (positive vs. negative lymph node metastasis)	ML	Logistic regression	Radiomic (textural) features	AUC = 0.77	Internal
Zhou H.	2020	China	2016–2018	26	NA	Cervical cancer	HISTOLOGY (benign vs. malignant)	ML, DL	*ML: Sparse nonnegative matrix factorization (SNMF) *DL: Deep belief network (DBN)	*ML input: Contrast-enhanced ultrasound image sequence; *DL input: Time-intensity curve (TIC)	* ML: NA *DL: accuracy = 0.86	Internal
Jin J.	2022	China	2010–2017	796	1102	Cervical cancer	METHODOLOGY (target automatic segmentation) METHODOLOGY (radiomics features reproducibility)	DL	U-Net with ResNet	2D gray-scale images	DSC = 0.90	Internal
Teng Y.	2022	China	2014–2018	148	8–15 (for each patient)	Cervical cancer	METHODOLOGY (target automatic segmentation) HISTOLOGY (positive vs. negative lymph node metastasis)	ML, DL	*DL: U-net with Resnet, attention Unet *ML: logistic regression	*DL input: 2D gray-scale images; *ML input: radiomics (shape, textural, wavelet) features	*DL auto-segmentation performance U-net with Resnet: DSC = 0.90; *Logistic regression prediction performance from features extracted from the ROI obtained from the DL attention U-Net: AUC = 0.75	Internal
Yi J.	2022	China	2014–2018	536	10–20 (for each patient)	Cervical cancer	HISTOLOGY (positive vs. negative lymph node metastasis)	ML	Support vector machine (SVM)	Radiomics (textural, wavelet) features	AUCs range among different scanners = [0.71–0.82]	Internal

TABLE 1 (Continued)

First author	Year	Country	Assessment time	Sample size	Number of images	Gynecological pathologies	Objective of the study	Type of AI used	AI specifics	Model input	Performance	Type of validation
							METHODOLOGY (radiomics features reproducibility)					
Chiappa V.	2021	Italy, Milan	2015–2020	70	NA	Miometrial tumor	HISTOLOGY (benign vs. malignant)	ML	Support vector machine (SVM) ensamble	Radiomic features (TRACE4©)	AUC = 0.86	Internal
Qin J.	2021	China	2012–2018	147	NA	Gestational trophoblastic neoplasia	PREDICTION OF RESISTANCE TO THERAPY	ML	Logistic regression	FIGO stage and Time-averaged mean velocity in UtA (UtA-TA mean)	Accuracy = 0.73	Internal
Fragomeni S.M.	2023	Italy	2017–2020	127	NA	Vulvar cancer	HISTOLOGY (metastatic lymph node detection)	ML	Morphonode predictive model (random forest classifiers)	Ultrasound features	AUC = 0.92	Internal

Note: Summary of the studies included in the paper listed according to different parameters (first author, year of publication, country, study assessment time, sample size, number of images collected, gynecological pathology, study objective) and AI models characteristics (type of AI used, AI specifics, model input, performance, validation type). In studies including multiple developed models, the results of the best performing model are reported in the "performance column." The performance has been reported in "AUC" when available, otherwise in "accuracy," then in "sensitivity/specificity" and then in other metrics which have been specified. The indicated performance refers to the external validation set, otherwise, to the internal validation set; if no validation was performed, this information has been reported.

Abbreviations: AUC, area under the receiver operating characteristic curve; DL, deep learning; DSC, dice similarity coefficient; ML, machine learning; NA, not specified; PLCO, Prostate, Lung, Colorectal and Ovarian database.

condition) and workflow (i.e., in most articles the time between the index test and the reference standard was reasonable) domains.

3.2 | Individual characteristics of the included studies

Most studies (35/50, 70.0%) were conducted as single center,[32–66] 7/50 (14.0%) were multicenter,[67–73] while 8/50 (16.0%) did not specify their design.[74–81] Single-center studies were conducted mostly in China (17/50, 34.0%),[33,43,44,46,48,49,51,52,54,57–63,65] followed by Italy (7/50, 14.0%).[36,40,41,47,56,66,73] Most papers were published between 2020 and 2023 (33/50, 66.0%).[36–54,57–66,70–73] The sample size ranged from 20 to 107,624 women; the majority of studies included less than 500 patients (32/50, 64.0%), while four studies included more than 1000 patients.

33/50 (66.0%) reported the diagnostic accuracy of ML models, 13/50 (26.0%) developed DL models, and 4/50 (8.0%) both ML and DL.

The primary aim of the large majority of the included studies was to predict histology (i.e., benign vs. malignant,[32–35,37–42,44–46,48–53,55,56,58–60,62–64,66–81]) while others focused on methodology (automatic segmentation,[43,54,61,62,70,72] reproducibility of radiomics features,[61,63] improvement of image quality[54]), prediction of therapy resistance,[65] progression to free survival (PFS)[47] and genetic mutation[36] (Table 1).

3.3 | Ovarian cancer

26/37 (70.3%) studies reported the performance of ML/DL models to discriminate between benign and malignant ovarian masses[32,34,35,37–42,45,48–50,52,67–69,71,74–81] while 4/37 (10.8%) explored the ability of such models to predict their histology (e.g., benign vs. borderline vs. malignant).[33,44,46,51] 2/37 (5.4%) focused on methodology (e.g., target automatic segmentation, improvement of image quality),[43,54] 1/37 (2.7%) assessed the performance of different models to predict PFS[47] and 1/37 (2.7%) BRCA mutation.[36]

Among those studies aimed at discriminating between benign and malignant adnexal masses, the study by Gao et al.[71] included the highest number of patients ($n = 107,624$). The authors developed a DL model from ultrasound images obtaining a performance similar to that of expert sonographers in the external validation set (AUC 0.87). Ben Van Calster et al.[67] developed an ML model including clinical and ultrasound variables of 1066 patients with AUC 0.95 in the internal validation set. Christiansen et al.[42] in a study including 758 patients developed ultrasound-based imaging DL model with AUC 0.95 in the internal validation test set. Hsu et al.[49] included 587 patients with adnexal masses and developed an ultrasound-based imaging DL model with accuracy 0.92 in the internal validation set (AUC not shown). Acharya et al.[81] in a study including 469 patients, developed an ultrasound imaging-based ML model with AUC 0.81 in the internal

validation set. Chen et al.[48] developed an ultrasound-based imaging DL model in a cohort of 420 patients with AUC 0.93 in the internal validation set. Ahmadi et al.[69] in a series of 305 patients, developed an ML model including clinical and ultrasound variables with AUC 0.91 in the internal validation set. Two studies[37,38] included data from the Prostate, Lung, Colorectal, and Ovarian database and both of them developed ML models. In one study, the authors included clinical and ultrasound variables with accuracy 0.81 (AUC not shown),[37] and in the other study they included ultrasound features with AUC 0.99 in the internal validation set.[38] The remaining studies[32,34,35,39–41,45,50,52,68,74–80] included a sample size of less than 300 patients or the sample size was not specified. No study, except for Gao[71] and Lucidame,[68] tested the model in an external validation set.

Among studies reporting the prediction of histology provided by different AI models, Qi et al.[44] built a combined clinical-radiomics model to discriminate between benign, borderline, and malignant serous ovarian tumors in a cohort of 265 patients and reported an AUC 0.91 in the internal validation set. From the same cohort, Wang et al.[46] tested the ability of a DL model based on ultrasound images to discriminate between benign, borderline, and malignant ovarian tumors with AUC 0.96 in the internal validation set. In the study of Wu et al.[33] the sample size was not specified (ultrasound images = 988), and the performance of the DL model was AUC 0.97 in the internal validation set. One study[51] included a sample size of less than 200 patients.

Three studies had more than one objective (first objective was to predict the specific histology and second objective was on methodology). Li et al.[72] in a series of 2021 patients, developed DL model for target automated segmentation and also realized DL model to discriminate between benign versus borderline versus malignant adnexal masses. For the first endpoint, the dice similarity coefficient was 0.92; for the second endpoint, the macro-F1 score was 0.75 in the external validation set (AUC not shown). In the other two studies, the sample size was low (35)[53] or not specified.[70]

Two studies built DL models to realize target automated segmentation,[43,54] and/or to improve image quality[54] with high performance.

Arezzo et al.[47] developed a ML model based on clinical and ultrasound variables to predict 12-month PFS in 64 patients with ovarian cancer, showing AUC 0.92 in the internal validation set.

Finally, Nero et al.[36] developed an ML model including radiomics features for predicting germline BRCA1/2 gene status in a cohort of 255 healthy patients showing accuracy 0.64 in the internal validation set (AUC not shown).

3.4 | Endometrial cancer

3/5 studies aimed at predicting malignancy, considering patients with abnormal uterine bleeding (one),[58] and regardless of symptoms (two)[55,56]; one study aimed to discriminate between low- and high-risk endometrial cancers[73] and one focused on the prediction of myometrial infiltration.[57] Ruan et al.[58] included the largest series of patients (1837) and developed a nomogram for the prediction of endometrial malignancy based on clinical and ultrasound variables showing AUC 0.91 in the internal validation set. In a cohort of 675 patients, Angioli et al.[56] developed a tool based on clinical and ultrasound variables to determine the probability of endometrial cancer (AUC 0.92 in the internal validation set). Michail et al.[55] developed an ML model to predict malignancy including a sample size less than 100 patients, but the model was not validated. Moro et al.[73] developed an ML model including clinical and ultrasound features to differentiate between high-risk endometrial cancers and the other three risk classes (low-, intermediate-, high-intermediate) with AUC 0.90 in the external validation set. Finally, Xu et al.[57] developed ML algorithms including clinical and ultrasound variables in the detection of deep myometrial invasion (sample size not specified) with accuracy 0.98 (AUC not shown) in the external validation set.

3.5 | Cervical cancer

Jin et al.[59] included a series of 172 patients and investigated the ability of noninvasive ultrasound-based radiomics methods in the preoperative discrimination between positive and negative lymph node metastasis in early cervical cancer, achieving AUC 0.77 in the internal validation set. In a subsequent study with a series of 148 patients, the same group compared the performance of automatic versus manual segmentation algorithms in lymph node metastasis detection by means of an ML algorithm based on radiomics features extracted from the segmented region of interest.[62] Models built with features based on DL automatic segmentation had higher performance than models built with features based on manual segmentation in the validation set (AUC 0.75 in the internal validation set). In a series of 536 patients, Yi et al.[63] developed an ML model including radiomics features to predict lymph node metastasis, and evaluated radiomics features reproducibility among different scanners concluding that the performance of the radiomics model is scanner-dependent (AUC range among different scanners 0.71–0.82).

In a series of 796 patients with cervical cancer, Jin et al.[61] developed an automated segmentation model showing a similar performance to that of manual segmentation (intraclass correlation coefficient 0.99).

Finally, Zhou et al.[60] conducted a study including 26 patients and developed an algorithm based on ultrasound contrast-enhanced images (time-intensity curve) to discriminate between malignant and benign cervix showing accuracy 0.86 (AUC not showed) in the internal validation set.

3.6 | Other cancers

Among studies concerning other gynecological tumors, Qin et al.[65] included a cohort of 147 patients affected by low-risk gestational trophoblastic neoplasia with myometrial invasion. Authors developed an ML model to predict methotrexate resistance. The model combined tumor vascularity with International Federation of Gynecology and Obstetrics prognostic scoring system, showing accuracy 0.73 in the

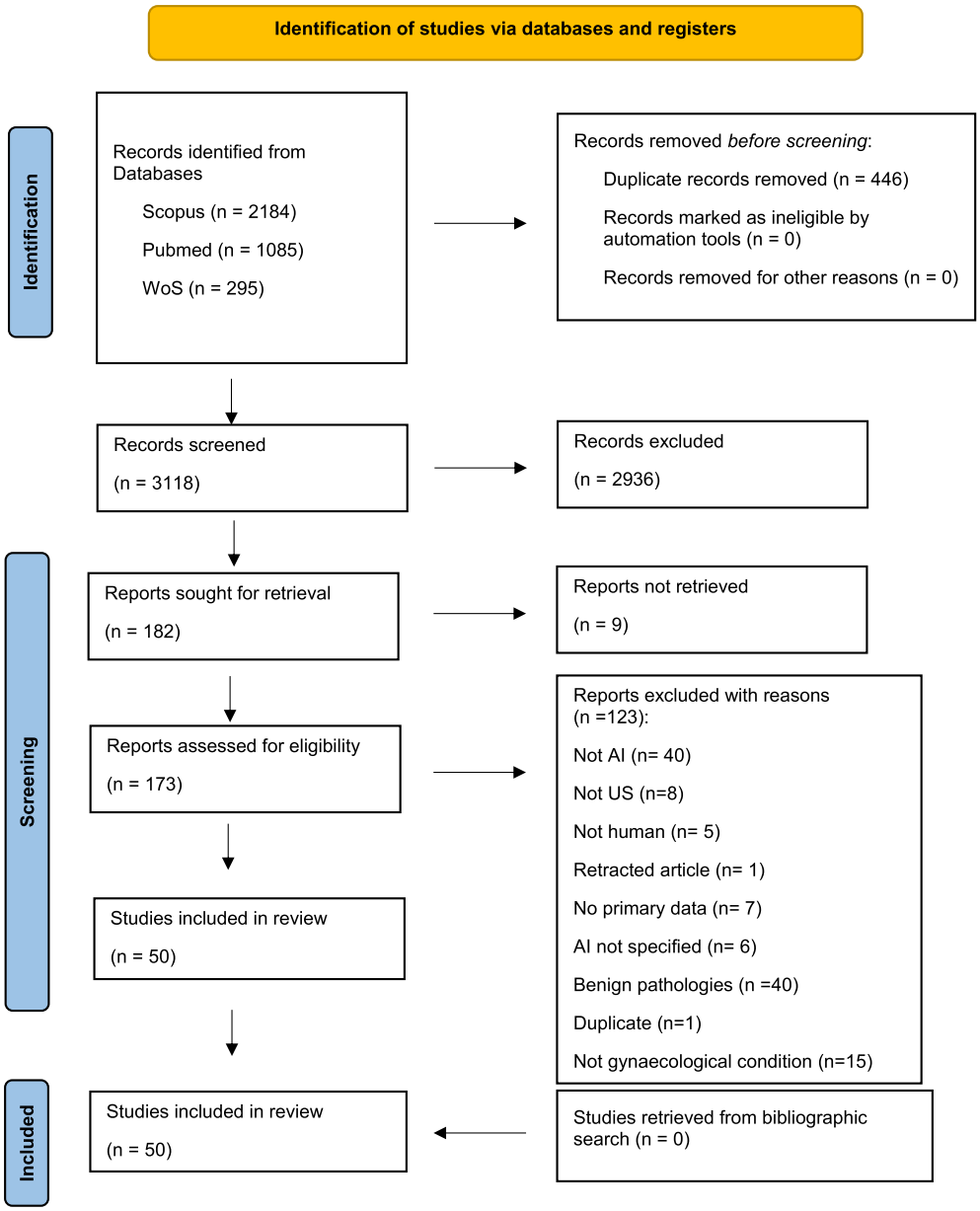

FIGURE 1 Flow diagram. Summary of the study identification and selection process, specifying whether papers were excluded or retrieved from bibliographic search with reasons appropriately clarified.

internal validation set (AUC not shown). In the study of Fragomeni et al.[66] including 127 patients, the authors developed an ML model (morphonode predictive model) aiming to discriminate between metastatic and nonmetastatic inguinal lymph nodes in vulvar cancer patients. The model showed AUC 0.92 in the internal validation set. Chiappa et al.[64] developed ML models including radiomics features to predict the risk of malignancy of uterine mesenchymal lesions in a series of 70 patients (AUC 0.86 in the internal validation set).

4 | DISCUSSION

The present review focuses on the role of AI applied to ultrasound imaging in gynecological oncology. Most articles reported the use of AI when applied to ovarian masses to define the diagnostic performance of ML in predicting histology. On the other hand, a limited number of studies on endometrial and cervical cancer have been published, focusing mostly on diagnostic performance of AI models in predicting pathological findings. The performance of ultrasound-based models was consistently high in most studies, demonstrating discriminative predictive ability and superiority when compared to non-AI methods. However, some methodological shortcomings should be mentioned: the external validation was presented only in few studies, the number of variables tested for modeling differed significantly among works, and the majority of studies were single-center including a low number of cases.

Our results agree with those reported in the published literature. Akazawa et al.[25] reported the role of AI in gynecological cancers and reviewed 71 articles (34 on cervical cancer, 21 on ovarian cancer, 13 on endometrial cancer, and three on uterine sarcoma). 35/71

studies used imaging data (i.e., MRI, CT, ultrasound, cytology, and hysteroscopy) including ultrasound images in only 5/35 studies. All five articles were on ovarian cancer. The authors highlighted the need to perform further studies in order to collect larger series of gynecological malignancies.

Shrestha et al,[26] included 61 articles and presented a similar scenario. Most studies were on MRI (35 articles) followed by CT (17 articles), positron emission tomography (6 articles), and ultrasound (8 articles). Again, most ultrasound image-based studies (7/8) were on ovarian cancer reporting that ML and DL models based on clinical, ultrasound, radiomics features or medical images may have great potential in supporting clinicians in the diagnosis and classification of ovarian tumors. Only one study included ultrasound images of cervical cancer to detect lymph node metastases using a DL method and showed satisfactory results compared with the radiologist's performance.

Ponsiglione et al.[28] in a review including 63 studies assessed the methodological rigor of radionics-based studies using imaging in the setting of ovarian cancer. Most articles were on CT and only 14 on ultrasound. Finally, Xu et al.[27] focused on AI performance in image-based ovarian cancer detection and the majority of studies (19/34) concerned ML and DL methods applied to ultrasound images (15/19 and 4/19, respectively), followed by MRI and CT. They concluded that AI algorithms excelled in the identification of ovarian cancer using medical radiography imaging, which manifested an equivalent or even better performance than independent detection by clinicians.

To the best of our knowledge, this is the first systematic review specifically dedicated to AI system performance applied to ultrasound in all fields of gynecological cancer. We conducted a comprehensive literature search in different databases to ensure the rigor of the study. We included data such as sample size, number of images, year of publication, geographical distribution, outcomes, as well as type of AI and families of variables included. Moreover, we assessed the quality of studies using the QUADAS-AI tool,[29] specifically adapted for AI research, which is a strength of this systematic review and will also guide future studies. However, we were unable to conduct metanalysis of the data, given variety in endpoint selection, validation, and performance metrics.

We believe that the present review can help readers to better understand the role of AI applied to ultrasound imaging. The AI could potentially impact our clinical management by improving the diagnostic accuracy and reducing time spent by the clinician which can be dedicated to relationships. The impact of AI on clinical management could be relevant. First, it could enhance diagnostic accuracy, thereby reducing the time spent by clinicians on diagnostic procedures. Second, it could free up time that could be allocated to more interpersonal aspects of care. Conversely, AI algorithms require the collection of large volumes of data to obtain extensive databases and they are created and managed by humans. In addition, the system needs a quality control process for data and a regular follow-up over time, demanding qualified and trained staff.

In conclusion, the main AI application to ultrasound in gynecology oncology regards improving preoperative diagnosis of ovarian masses to help clinicians and surgeons plan the best treatments for patients also when expert ultrasound examiners are not available. AI applications are still lacking for other pathologies including myometrial lesions, endometrial and cervical cancers, as well as to predict tumor response to therapy, genetic mutation status, and disease-free survival. AI and radiomics applied to ultrasound, which is widely available in clinical settings, can open up further research and new strategies in the management of gynecological oncology patients. For example, AI may help to predict histological factors and molecular profile preoperatively in order to better personalize treatment (i.e., POLE mutation in endometrial cancer, PDL-1 in cervical cancer, LVSI in early cervical cancer). It may also have other applications in predicting treatment response after chemotherapy and recurrence in ovarian and cervical cancers.

AUTHOR CONTRIBUTIONS

Francesca Moro: Conceptualization; data curation; investigation; project administration; supervision; visualization; writing – review and editing. **Marianna Ciancia:** Data curation; investigation; writing – original draft. **Drieda Zace:** Methodology; writing – original draft. **Marica Vagni:** Data curation; formal analysis; investigation; visualization; writing – original draft. **Huong Elena Tran:** Data curation; formal analysis; investigation; visualization; writing – original draft. **Maria Teresa Giudice:** Data curation; investigation. **Sofia Gambigliani Zoccoli:** Data curation; investigation. **Floriana Mascilini:** Data curation. **Francesca Ciccarone:** Data curation. **Luca Boldrini:** Supervision. **Francesco D'Antonio:** Supervision; writing – review and editing. **Giovanni Scambia:** Supervision. **Antonia Carla Testa:** Supervision; writing – review and editing.

CONFLICT OF INTEREST STATEMENT

The authors declare that there are no conflicts of interest.

DATA AVAILABILITY STATEMENT

Only publicly available data were used in this study, and data sources and handling of these data are described in the Materials and Methods. QUADAS-AI tool is used for the quality assessment of the included studies. Further information is available from the corresponding author upon request.

ETHICS STATEMENT

The present systematic review was conducted and reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement. The protocol was published in PROSPERO (registration record CRD42023427088).

ORCID

Francesca Moro https://orcid.org/0000-0002-5070-7245
Giovanni Scambia https://orcid.org/0000-0002-9503-9041

REFERENCES

1. Fischerova D, Burgetova A. Imaging techniques for the evaluation of ovarian cancer. *Best Pract Res Clin Obstet Gynaecol*. 2014;28:697-720.
2. Testa AC, Di Legge A, Virgilio B, et al. Which imaging technique should we use in the follow up of gynaecological cancer? *Best Pract Res Clin Obstet Gynaecol*. 2014;28:769-791.

3. Testa AC, Di Legge A, De Blasis I, et al. Imaging techniques for the evaluation of cervical cancer. *Best Pract Res Clin Obstet Gynaecol*. 2014;28:741-768.
4. Epstein E, Fischerova D, Valentin L, et al. Ultrasound characteristics of endometrial cancer as defined by international endometrial tumor analysis (IETA) consensus nomenclature: prospective multicenter study. *Ultrasound Obstet Gynecol*. 2018;51:818-828.
5. Fischerová D, Cibula D. The role of ultrasound in primary workup of cervical cancer staging (ESGO, ESTRO, ESP cervical cancer guidelines). *Ceska Gynekol*. 2019;84:40-48.
6. Moro F, Esposito R, Landolfo C, et al. Ultrasound evaluation of ovarian masses and assessment of the extension of ovarian malignancy. *Br J Radiol*. 2021;94:20201375.
7. Moruzzi MC, Bolomini G, Esposito R, et al. Diagnostic performance of ultrasound in assessing the extension of disease in advanced ovarian cancer. *Am J Obstet Gynecol*. 2022;227:601.e1-601.e20.
8. Timmerman D, Testa AC, Bourne T, et al. International ovarian tumor analysis group. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the international ovarian tumor analysis group. *J Clin Oncol*. 2005;23:8794-8801.
9. Timmerman D, Testa AC, Bourne T, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. *Ultrasound Obstet Gynecol*. 2008;31:681-690.
10. Ameye L, Timmerman D, Valentin L, et al. Clinically oriented three-step strategy for assessment of adnexal pathology. *Ultrasound Obstet Gynecol*. 2012;40:582-591.
11. van Calster B, van Hoorde K, Valentin L, et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. *BMJ*. 2014;349:g5920.
12. Landolfo C, Bourne T, Froyman W, et al. Benign descriptors and ADNEX in two-step strategy to estimate risk of malignancy in ovarian tumors: retrospective validation in IOTA5 multicenter cohort. *Ultrasound Obstet Gynecol*. 2023;61:231-242.
13. Timmerman D, Planchamp F, Bourne T, et al. ESGO/ISUOG/IOTA/ESGE consensus statement on preoperative diagnosis of ovarian tumours. *Facts Views Vis Obgyn*. 2021;13:107-130.
14. Concin N, Matias-Guiu X, Vergote I, et al. ESGO/ESTRO/ESP guidelines for the management of patients with endometrial carcinoma. *Int J Gynecol Cancer*. 2021;31:12-39.
15. Cibula D, Raspollini MR, Planchamp F, et al. ESGO/ESTRO/ESP guidelines for the management of patients with cervical cancer—update 2023. *Int J Gynecol Cancer*. 2023;33:649-666.
16. Bi WL, Hosny A, Schabath MB, et al. Artificial intelligence in cancer imaging: clinical challenges and applications. *CA Cancer J Clin*. 2019;69:127-157.
17. Lecointre L, Dana J, Lodi M, Akladios C, Gallix B. Artificial intelligence-based radiomics models in endometrial cancer: a systematic review. *Eur J Surg Oncol*. 2021;47:2734-2741.
18. Rizzo S, Manganaro L, Dolciami M, Gasparri ML, Papadia A, del Grande F. Computed tomography based radiomics as a predictor of survival in ovarian cancer patients: a systematic review. *Cancers*. 2021;13:1-11.
19. Smith RG, Eckroth J. Building AI applications: yesterday, today, and tomorrow. *AI Mag*. 2017;38:6-22.
20. Bishop Christopher M. *Pattern Recognition and Machine Learning*. Springer; 2006.
21. Xin Y, Kong L, Liu Z, et al. Machine learning and deep learning methods for cybersecurity. *IEEE Access*. 2018;6:35365-35381.
22. Hatt M, Parmar C, Qi J, El Naqa I. Machine (deep) learning methods for image processing and radiomics. *IEEE Trans Radiat Plasma Med Sci*. 2019;3:104-108.
23. Rogers W, Seetha T, Refaee S, Lieverse TIY, Granzier RWY, Ibrahim R. Radiomics: from qualitative to quantitative imaging. 2020.
24. Lambin P, Leijenaar RTH, Deist TM, et al. Radiomics: the bridge between medical imaging and personalized medicine. *Nat Rev Clin Oncol*. 2017;14:749-762.
25. Akazawa M, Hashimoto K. Artificial intelligence in gynecologic cancers: current status and future challenges—a systematic review. *Artif Intell Med*. 2021;120:102164.
26. Shrestha P, Poudyal B, Yadollahi S, 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;166:596-605.
27. Xu H-L, Gong T-T, Liu F-H, et al. Artificial intelligence performance in image-based ovarian cancer identification: a systematic review and meta-analysis. *EClinicalMedicine*. 2022;53:101662.
28. Ponsiglione A, Stanzione A, Spadarella G, et al. Ovarian imaging radiomics quality score assessment: an EuSoMII radiomics auditing group initiative. *Eur Radiol*. 2023;33:2239-2247.
29. Sounderajah V, Ashrafian H, Rose S, et al. A quality assessment tool for artificial intelligence-centered diagnostic test accuracy studies: QUADAS-AI. *Nat Med*. 2021;27:1663-1665.
30. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155:529-536.
31. Yang B, Mallett S, Takwoingi Y, et al. QUADAS-C: a tool for assessing risk of bias in comparative diagnostic accuracy studies. *Ann Intern Med*. 2021;174:1592-1599.
32. Aramendía-Vidaurreta V, Cabeza R, Villanueva A, Navallas J, Alcázar JL. Ultrasound image discrimination between benign and malignant adnexal masses based on a neural network approach. *Ultrasound Med Biol*. 2016;42:742-752.
33. Wu C, Wang Y, Wang F. Deep learning for ovarian tumor classification with ultrasound images. *Advances in Multimedia Information Processing—PCM 2018*. Springer; 2018:395-406.
34. Martínez-Más J, Bueno-Crespo A, Khazendar S, et al. Evaluation of machine learning methods with Fourier transform features for classifying ovarian tumors based on ultrasound images. *PLoS One*. 2019;14:e0219388.
35. Stukan M, Badocha M, Ratajczak K. Development and validation of a model that includes two ultrasound parameters and the plasma D-dimer level for predicting malignancy in adnexal masses: an observational study. *BMC Cancer*. 2019;19:19.
36. Nero C, Ciccarone F, Boldrini L, et al. Germline BRCA 1-2 status prediction through ovarian ultrasound images radiogenomics: a hypothesis generating study (PROBE study). *Sci Rep*. 2020;10:16511.
37. Akter L, Akhter N. Detection of ovarian malignancy from combination of CA125 in blood and TVUS using machine learning. *International Conference on Trends in Computational and Cognitive Engineering*. Springer; 2021:279-289.
38. Akter L, Akhter N. Ovarian cancer prediction from ovarian cysts based on TVUS using machine learning algorithms. *International Conference on Big Data, IoT, and Machine Learning*. Springer; 2021:51-61.
39. Al-karawi D, Al-Assam H, Du H, et al. An evaluation of the effectiveness of image-based texture features extracted from static B-mode ultrasound images in distinguishing between benign and malignant ovarian masses. *Ultrason Imaging*. 2021;43:124-138.
40. Chiappa V, Interlenghi M, Bogani G, et al. A decision support system based on radiomics and machine learning to predict the risk of malignancy of ovarian masses from transvaginal ultrasonography and serum CA-125. *Eur Radiol Exp*. 2021;28:5.
41. Chiappa V, Bogani G, Interlenghi M, et al. The adoption of radiomics and machine learning improves the diagnostic processes of women with ovarian MAsses (the AROMA pilot study). *J Ultrasound*. 2021;24:429-437.
42. Christiansen F, Epstein EL, Smedberg E, Åkerlund M, Smith K, Epstein E. Ultrasound image analysis using deep neural networks for

discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment. *Ultrasound Obstet Gynecol*. 2021;57:155-163.
43. Jin J, Zhu H, Zhang J, et al. Multiple U-Net-based automatic segmentations and radiomics feature stability on ultrasound images for patients with ovarian cancer. *Front Oncologia*. 2021;10:614201.
44. Qi L, Chen D, Li C, et al. Diagnosis of ovarian neoplasms using nomogram in combination with ultrasound image-based radiomics signature and clinical factors. *Front Genet*. 2021;12:753948.
45. Ștefan PA, Lupean RA, Mihu CM, et al. Ultrasonography in the diagnosis of adnexal lesions: the role of texture analysis. *Diagnostics*. 2021;11:812.
46. Wang H, Liu C, Zhao Z, et al. Application of deep convolutional neural networks for discriminating benign, borderline, and malignant serous ovarian tumors from ultrasound images. *Front Oncol*. 2021;11:770683.
47. Arezzo F, Cormio G, La Forgia D, et al. A machine learning approach applied to gynecological ultrasound to predict progression-free survival in ovarian cancer patients. *Arch Gynecol Obstet*. 2022;306:2143-2154.
48. Chen H, Yang BW, Qian L, et al. Deep learning prediction of ovarian malignancy at US compared with O-RADS and expert assessment. *Radiology*. 2022;304:106-113.
49. Hsu ST, Su YJ, Hung CH, Chen MJ, Lu CH, Kuo CE. Automatic ovarian tumors recognition system based on ensemble convolutional neural network with ultrasound imaging. *BMC Med Inform Decis Mak*. 2022;22:298.
50. Srilatha K, Jayasudha FV, Sumathi M, Chitra P. Automated ultrasound ovarian tumour segmentation and classification based on deep learning techniques. *Advances in Electrical and Computer Technologies*. Springer; 2022:59-70.
51. Tang ZP, Ma Z, He Y, et al. Ultrasound-based radiomics for predicting different pathological subtypes of epithelial ovarian cancer before surgery. *BMC Med Imaging*. 2022;22:22.
52. Yin Q, Zhong M, Wang Z, Sheng XJ. Clinical analysis of 137 cases of ovarian tumors in pregnancy. *J Oncol*. 2022;2022:1-9.
53. Zou Y, Amidi E, Luo H, Zhu Q. Ultrasound-enhanced U-Net model for quantitative photoacoustic tomography of ovarian lesions. *Photoacoustics*. 2022;28:28.
54. Chen L, Qiao C, Wu M, et al. Improving the segmentation accuracy of ovarian-tumor ultrasound images using image Inpainting. *Bioengineering*. 2023;10:184.
55. Michail G, Karahaliou A, Skiadopoulos S, et al. Texture analysis of perimenopausal and post-menopausal endometrial tissue in grayscale transvaginal ultrasonography. *Br J Radiol*. 2007;80:609-616.
56. Angioli R, Capriglione S, Aloisi A, et al. REM (risk of endometrial malignancy): a proposal for a new scoring system to evaluate risk of endometrial malignancy. *Clin Cancer Res*. 2013;19:5733-5739.
57. Xu J, Zeng H, He S, Qin L. Comparison of machine learning algorithms in the context of deep diagnosis of endometrial carcinoma with myometrial invasion. *Proceedings—2022 9th International Conference on Digital Home, ICDH 2022*. Institute of Electrical and Electronics Engineers Inc.; 2022:100-106.
58. Ruan H, Chen S, Li J, et al. Development and validation of a nomogram prediction model for endometrial malignancy in patients with abnormal uterine bleeding. *Yonsei Med J*. 2023;64:197-203.
59. Jin X, Ai Y, Zhang J, et al. Noninvasive prediction of lymph node status for patients with early-stage cervical cancer based on radiomics features from ultrasound images. *Eur Radiol*. 2020;30:4117-4124.
60. Zhou H, Wang S, Zhang T, Liu D, Yang K. Ultrasound image analysis technology under deep belief networks in evaluation on the effects of diagnosis and chemotherapy of cervical cancer. *J Supercomput*. 2021;77:4151-4171.
61. Jin J, Zhu H, Teng Y, Ai Y, Xie C, Jin X. The accuracy and radiomics feature effects of multiple U-Net-based automatic segmentation models for transvaginal ultrasound images of cervical cancer. *J Digit Imaging*. 2022;35:983-992.
62. Teng Y, Ai Y, Liang T, et al. The effects of automatic segmentations on preoperative lymph node status prediction models with ultrasound radiomics for patients with early stage cervical cancer. *Technol Cancer Res Treat*. 2022;21:153303382210993.
63. Yi J, Lei X, Zhang L, et al. The influence of different ultrasonic machines on Radiomics models in prediction lymph node metastasis for patients with cervical cancer. *Technol Cancer Res Treat*. 2022;21:15330338221118412.
64. Chiappa V, Interlenghi M, Salvatore C, et al. Using rADioMIcs and machine learning with ultrasonography for the differential diagnosis of myometRiAL tumors (the ADMIRAL pilot study). Radiomics and differential diagnosis of myometrial tumors. *Gynecol Oncol*. 2021;161:838-844.
65. Qin J, Zhang S, Poon L, et al. Doppler-based predictive model for methotrexate resistance in low-risk gestational trophoblastic neoplasia with myometrial invasion: prospective study of 147 patients. *Ultrasound Obstet Gynecol*. 2021;57:829-839.
66. Fragomeni SM, Moro F, Palluzzi F, et al. Evaluating the risk of inguinal lymph node metastases before surgery using the morphonode predictive model: a prospective diagnostic study in vulvar cancer patients. *Cancers*. 2023;15:1121.
67. van Calster B, Timmerman D, Lu C, et al. Preoperative diagnosis of ovarian tumors using Bayesian kernel-based methods. *Ultrasound Obstet Gynecol*. 2007;29:496-504.
68. Lucidarme O, Akakpo JP, Granberg S, et al. A new computer-aided diagnostic tool for non-invasive characterisation of malignant ovarian masses: results of a multicentre validation study. *Eur Radiol*. 2010;20:1822-1830.
69. Ahmadi E, Javadi H, Khansefid A, Asadi A, Ebadzadeh MM, Timmerman D. Fuzzy decision tree learning for preoperative classification of adnexal masses. In: *HEALTHINF 2011—Proceedings of the International Conference on Health Informatics*. SciTePress; 2011:364-375.
70. Hussein IJ, Burhanuddin MA, Mohammed MA, et al. Fully automatic segmentation of gynaecological abnormality using a new Viola–Jones model. *Comput Mater Cont*. 2021;66:3161-3182.
71. Gao Y, Zeng S, Xu X, et al. Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in china: a retrospective, multicentre, diagnostic study [Internet]. 2022. 179–87 Available from: www.thelancet.com/
72. Li J, Chen Y, Zhang M, et al. A deep learning model system for diagnosis and management of adnexal masses. *Cancers*. 2022;14:5291.
73. Moro F, Albanese M, Boldrini L, et al. Developing and validating ultrasound-based radiomics models for predicting high-risk endometrial cancer. *Ultrasound Obstet Gynecol*. 2022;60:256-268.
74. Acharya UR, Sree VS, Saba L, Molinari F, Guerriero S, Suri JS. Ovarian tumor characterization and classification: a class of GyneScan™ systems. *Annu Int Conf IEEE Eng Med Biol Soc*. 2012;2012:4446-4449.
75. Acharya UR, Vinitha Sree S, Muthu Rama Krishnan M, et al. Ovarian tumor characterization using 3D ultrasound. *Technol Cancer Res Treat*. 2012;11:543-552.
76. Faschingbauer F, Beckmann MW, Weyert Goecke T, et al. Automatic texture-based analysis in ultrasound imaging of ovarian masses. *Ultraschall Medizin*. 2013;34:145-150.
77. Acharya UR, Sree VS, Saba L, Molinari F, Guerriero S, Suri JS. Ovarian tumor characterization and classification using ultrasound—a new online paradigm. *J Digit Imaging*. 2013;26:544-553.
78. Acharya UR, Mookiah MRK, Vinitha Sree S, et al. Evolutionary algorithm-based classifier parameter tuning for automatic ovarian cancer tissue characterization and classification. *Ultraschall Medizin*. 2014;35:237-245.

79. Acharya UR, Sree VS, Kulshreshtha S, et al. GyneScan: an improved online paradigm for screening of ovarian cancer via tissue characterization. *Technol Cancer Res Treat*. 2014;13:529-540.
80. Pathak H, Kulkarni V. Identification of ovarian mass through ultrasound images using machine learning techniques. *2015 IEEE International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*. IEEE; 2015: 137-140.
81. Acharya UR, Akter A, Chowriappa P, et al. Use of nonlinear features for automated characterization of suspicious ovarian tumors using ultrasound images in fuzzy Forest framework. *Int J Fuzzy Syst*. 2018; 20:1385-1402.

SUPPORTING INFORMATION

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

How to cite this article: Moro F, Ciancia M, Zace D, et al. Role of artificial intelligence applied to ultrasound in gynecology oncology: A systematic review. *Int J Cancer*. 2024;155(10): 1832-1845. doi:10.1002/ijc.35092