Seminars in Cancer Biology 96 (2023) 82–99

Contents lists available at ScienceDirect

Seminars in Cancer Biology

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

Artificial intelligence-based risk stratification, accurate diagnosis and treatment prediction in gynecologic oncology

Yuting Jiang [a,b,1], Chengdi Wang [a,b,1], Shengtao Zhou [a,b,*]

[a] Department of Obstetrics and Gynecology, Key Laboratory of Birth Defects and Related Diseases of Women and Children of MOE and State Key Laboratory of Biotherapy, West China Second Hospital, Sichuan University and Collaborative Innovation Center, Chengdu, Sichuan 610041, China
[b] Department of Pulmonary and Critical Care Medicine, State Key Laboratory of Respiratory Health and Multimorbidity, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

ARTICLE INFO

Keywords:
Artificial intelligence
Cancer imaging
Gynecologic cancer
Molecular profiling
Precision oncology

ABSTRACT

As data-driven science, artificial intelligence (AI) has paved a promising path toward an evolving health system teeming with thrilling opportunities for precision oncology. Notwithstanding the tremendous success of oncological AI in such fields as lung carcinoma, breast tumor and brain malignancy, less attention has been devoted to investigating the influence of AI on gynecologic oncology. Hereby, this review sheds light on the ever-increasing contribution of state-of-the-art AI techniques to the refined risk stratification and whole-course management of patients with gynecologic tumors, in particular, cervical, ovarian and endometrial cancer, centering on information and features extracted from clinical data (electronic health records), cancer imaging including radiological imaging, colposcopic images, cytological and histopathological digital images, and molecular profiling (genomics, transcriptomics, metabolomics and so forth). However, there are still noteworthy challenges beyond performance validation. Thus, this work further describes the limitations and challenges faced in the real-word implementation of AI models, as well as potential solutions to address these issues.

Abbreviations: AI, artificial intelligence; ANN, artificial neuronal network; AUC, area under curve; AUROC, area under the receiving operator characteristic curve; CART, classification and regression trees; CA125, cancer antigen 125; CA19–9, carbohydrate 19–9; cfDNA, cell-free DNA; C-index, concordance index; CIN, cervical intraepithelial neoplasia; CNN, convolutional network; CNV, copy-number variation; CPTAC, clinical proteomic tumor analysis consortium; CRP, C-reactive protein; CT, computed tomography; ctDNA, circulating tumor DNA; CUP, cancer of unknown primary; CV, cross-validation; DCNN, deep convolutional neural network; DFI, disease-free interval; DFS, disease-free survival; DL, deep learning; DLS, deep learning system; DMI, deep myometrial invasion; dMMR, different mismatch repair; DNN, deep neural network; DS, dual stain; DSC, Dice similarity coefficient; DSS, disease-specific survival; EHR, electronic health record; EOC, epithelial ovarian cancer; FDA, Food and Drug Administration; FIGO, the International Federation of Gynecology and Obstetrics; HER2, human epidermal growth factor receptor 2; HGSOC, high-grade serous ovarian cancer; HIC, high-income country; HIPAA, the Health Insurance Portability and Accountability Act; HPV, human papillomavirus; HRD, homologous recombination deficiency; H&E, hematoxylin and eosin-stained; KNN, k-nearest neighbor; LASSO, least absolute shrinkage and selection operator; LDH, lactate dehydrogenase; LMIC, low- and middle-income country; LNM, lymph node metastasis; LVSI, lymphovascular space invasion; ML, machine learning; MRI, magnetic resonance imaging; MSI, microsatellite instability; NACT, neoadjuvant chemotherapy; NC, not calculated; NLP, natural language processing; OS, overall survival; O-RADS, the Ovarian-Adnexal Reporting and Data System; PARP, poly-ADP ribose polymerase; PET, positron emission tomography; PFS, Progression-free survival; ROC, recurrent ovarian cancer; RCT, randomized clinical trial; SCS, secondary cytoreductive surgery; scRNA, single cell RNA; SEER, Surveillance, Epidemiology, and End Results; SHG, second-harmonic generation; SVM, support vector machine; TBS, the Bethesda system; TCGA, The Cancer Genome Atlas; THG, third-harmonic generation; TME, tumor microenvironment; US, ultrasonography; VEGF, vascular endothelial growth factor; VIA, visual inspection with acetic acid; WGS, whole-genome sequencing; WHO, World Health Organization; WSI, whole-slide images; xAI, explainable AI; ^{18}F-FDG PET/CT, ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography.

* Correspondence to: Department of Obstetrics and Gynecology, Key Laboratory of Birth Defects and Related Diseases of Women and Children of MOE and State Key Laboratory of Biotherapy, West China Second Hospital, Sichuan University and Collaborative Innovation Center, Chengdu 610041, China.
E-mail address: shengtaozhou@scu.edu.cn (S. Zhou).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.semcancer.2023.09.005
Received 17 December 2022; Received in revised form 27 August 2023; Accepted 25 September 2023
Available online 30 September 2023
1044-579X/

1. Introduction

Artificial intelligence (AI) is a vast domain that encompasses computer science and statistics and refers to the ability of a machine to execute cognitive functions imitating human intelligence [1,2]. The last decade has witnessed an impressive resurgence in artificial intelligence, particularly deep learning (DL) for clinical decision-making, and predominantly propelled by the rapid advances in computational power and the availability to exponentially growing medical data [3]. Oncology is especially ripe for transformational changes yielded by AI, with promising preliminary results in automated Gleason grading of prostate biopsies [4,5], dermatologist-level classification of skin diseases and so forth [6–9].

Globally, gynecologic cancers account for roughly 15% of total new cases and total new deaths among female cancer patients in 2020 [10]. Cancer of the cervix uteri (7.7%) is the most lethal gynecologic cancer, followed by ovary (4.7%), corpus uteri (2.2%), vulva and vagina worldwide [10]. Considering the terrific performance in clinically relevant tasks, AI is poised to widely reshape the gynecologic cancer care intended for enhanced quality of life as well. As an example, cervical cancer is a highly preventable and curable disease when detected early and treated timely, and proven and cost-effective cervical cancer screening programs mostly rely on Papanicolaou smear cytology in combination with HPV detection, visual inspection with acetic acid (VIA), colposcopy, etcetera. Herein, in 2020, the Director-General of World Health Organization (WHO) launched strategy to accelerate the global elimination of cervical cancer as a public health problem in 194 countries [11]. Nevertheless, few cancers reflect unfairness as much as cervical cancer, evidenced by that in low- and middle-income countries (LMICs) the mortality rate is nearly three times as high and the incidence is approximately twice as high as in high-income countries (HICs) [10,12,13]. Accordingly, it seems attractive and practical to apply AI-based methods in resource-constrained settings where for instance, access to microscopy diagnostics is inadequate and the qualities of such visual inspection vary greatly. Noticeably, digital microscopy supported by a cloud-based deep learning system (DLS) was successfully implemented in rural Kenya and was utilized to detect squamous cell atypia in Papanicolaou tests with a high sensitivity (95.7%) [14]. In addition, AI has the potential to address the difficulty in early detection of ovarian cancer using non-invasive test [15] and to predict the molecular-based classification of endometrial carcinoma from histopathology images [16], both of which are of paramount importance for risk stratification and personalized management [17,18].

Over the course of diagnosis, treatment and surveillance, a cornucopia of muti-dimensional data will be generated for an individual with gynecologic malignancy, spanning clinical presentation, patient history, radiology, histology, laboratory tests and genomics, and each modality may contribute to distinct insights on the state of gynecologic oncology. However, it is often too complex to subjectively analyze the latent relations and patterns across various data modalities, partially even exceeding the ability of human [19]. For this reason, a remarkable opportunity for the clinical implementation of AI emerges to aggregate and incorporate the complementary digital assets and clinical context from large-scale populations to identify novel multimodal cancer biomarkers, which probably could function as surrogates for current yet high-priced markers. On the basis of primary use, AI-enabled biomarkers for gynecologic cancer can be divided into diagnostic [20–22], prognostic or predictive markers of response and resistance to therapy in the form of qualitative or quantitative ways [23,24], informing better personalized care for patients with gynecologic cancers [25].

Herein, this review highlights the recent AI-based attempts and ever-growing contributions to the management of gynecologic oncology, which have bridged the development-to-implementation gap (Table 1). Moreover, this work outlines future directions on AI application in precision oncology and maps the key challenges of AI-driven healthcare.

2. Understanding AI tools and techniques

The concept of AI, coined in the 1956, was originally referred to as "thinking machine" [9]. AI is a broad field comprising machine learning and deep learning [26], and further categorized into supervised, weakly supervised, unsupervised learning and reinforcement learning (Fig. 1).

Two well-known examples of supervised strategies are hand-crafted methods and representation learning methods. The former is typically performed with machine learning algorithms, which enable machines to process input data, learn and then elaborate predictions without explicit pre-programming [27]. As the dominant form in medical AI, supervised machine learning uses explicit data with "ground truth'' annotations to optimize a model, in which unlabeled input will be classified into pre-defined categories [28].

To alleviate the heavy labeling burden of the supervised learning, weakly supervised and unsupervised learning are gaining attention. As a subfield of supervised learning, weakly supervised learning involves multiple-instance learning, graph convolutional networks and vision transformers. Weak supervision is particularly suitable when the data labeling is incomplete, inaccurate or inadequate to construct a model with excellent performance [29]. Unsupervised learning is capable of investigating intrinsic patterns and subtypes in dataset dispensing with labels, and is classified into fully unsupervised and self-supervised methods. In a sense, unsupervised learning algorithms are allowed to accomplish more complex tasks than that the supervised learning methods can handle with [30]. As the science of decision making, reinforcement learning, permitting interactive feedback, is optimal to solve complex, sequential problems [31], and robotic-assisted surgery may gain benefits from reinforcement learning [32–34].

But the capability to deal with natural raw data is limited for conventional ML algorithms. Instead, deep learning, a newer sub-category of machine learning, is able to automatically learn representations of data by virtue of artificial neuronal network structure (Fig. 2) [3]. The most important advantage of DL approaches over other ML is the competence to execute feature engineering by itself from limited series of features provided in the training dataset, and thereby may be utilized to mitigate the shortage of experienced physicians [35]. In addition, empowered by deep learning, natural language processing (NLP) is striving to capture linguistic nuances and to comprehend textual entailment, hopeful to boost EHR-based cancer research [36].

3. Early detection of gynecologic oncology

Early cancer detection involves the input and estimate of data at different time point to determine whether or not the individuals are in need of supplementary examination and assessment, aimed to optimize the survivorship of patients with gynecologic cancers (Fig. 3) [9]. For gynecologic oncology, the data streams are mainly in the form of visual inspection, medical imaging such as ultrasonography and colposcopic images, and serological markers, such as cancer antigen 125 (CA125) [37–39].

3.1. Cervical cancer: a preventable disease

In comparison with ovarian cancer and uterine malignancy, cervical cancer is regarded as a potentially completely preventable disease since it may take several years for cervical intraepithelial neoplasia (CIN) to progress to invasive cancer, offering opportunities for early detection and timely intervention [40]. Typically, the screening workflow of cervical cancer moves in a stepwise fashion from cytology combined with HPV testing, to colposcopy and biopsy [41]. Hence, a field in which the adoption of AI is pretty appealing is the early diagnosis of cervical precancerous lesions and tumors.

3.1.1. HPV testing

Given the integration of high-risk human papillomavirus (hrHPV)

Table 1
Representative studies in AI for gynecologic oncology.

Year	Task	Modality	Model type	Study design	Size of dataset	Against Human	Performance	Ref.
Cervical Cancer								
	Screening I	**HPV testing**						
2019	HPV	Microholography	DL	-	Training: 12 images; Test: 28 patients	No	Concordance: 100%	[44]
2020	HPV integration sites	Gene expressing	DL	Retrospective	Training: 3608 integration sites; Validation: 584 integration sites Test: 4662 integration sites	No	AUC: 0.7175	[43]
	Screening II	**Cytology**						
2021	TBS classification	Cytopathology	DL	-	Training: 350 patients; Validation: 390 patients; Test: 361 patients	Yes	AUC: 0.94; Sensitivity: 95.7%; Specificity: 84.7%	[14]
2021	TBS classification	Cytopathology	Dual-path DCNN	Retrospective	Training: 13486 WSIs; Validation: 2486 WSIs; Test: 3331 WSIs	No	AUC: 0.925	[45]
2021	TBS classification	Cytopathology	DL	Retrospective	3545 WSIs for training and validation; Test: 1170 WSIs	Yes	AUC: 0.979; Sensitivity: 95.1%; Specificity: 93.5%	[46]
2021	Abnormal cervical cells	Cytopathology	CNN	Retrospective	7030 images for training and internal validation; External validation: 35013 images; Test: 3970 images	Yes	AUC: 0.99 Accuracy: 98.0% (human: 78.5%); Sensitivity: 98.0% (human: 63.2%); Specificity: 98.0% (human: 93.8%)	[47]
2021	TBS classification	Cytopathology	DL;ML	Retrospective Prospective	Training: 81727 smears; Validation: 10-fold CV Test: 34403 smears	Yes	Retrospective: Accuracy: 75.24% (senior expert: 80.16%); Prospective: Sensitivity: 94.7%	[48]
2021	Detection of CIN3 +	Cytopathology (p16/ki-67)	DL	Retrospective	Training: 345 patients; Validation: 86 patients; Test: 3095 patients	Yes	AUC: 0.74; Sensitivity: 87%; (human: 87%); Specificity: 45.6% (human: 40.5%);	[50]
	Screening III	**Colposcopy**						
2019	Detection of CIN 2 +	Cervico-graphy	DL	Retrospective	Training: 744 patients; Validation: 324 patients; Test: 8259 patients	No	AUC: 0.91	[57]
2020	Colposcopic grading; Guiding biopsies	Colposcopic image	DL	Retrospective	Training: 13604 patients; Validation: 1944 patients; Test: 3887 patients	Yes	Grading colposcopic impressions: Agreement: 82.2% (human: 65.9%); Kappa: 0.75 (human: 0.516); Guiding biopsies: mean-intersection-over-union: 0.758	[52]
2020	Non-cancer vs CIN vs cancer	Time-lapsed colposcopic image	DL	Retrospective	Training: 6133 patients; Validation:1535 patients	Yes	Accuracy: 78.33% (human: 75.11%)	[53]
2020	Non-cancer vs CIN vs cancer	Real-time noninvasive imaging	DL; ML	Retrospective	a. Training: 220 images; Validation: 76 images; Test: 41 images; b. Training: 96 images; Validation: 32 images; Test: 36 patients	No	AUC $\geq$ 0.92; Kappa: 0.86	[59]
2021	CIN1 vs CIN2 +	Colposcopic image	Contra-stive learning	Retrospective	Training: 6023 patients; Validation: 431 patients; Test: 2150 patients	Yes	Accuracy: 61.84% (human: 58.68%)	[54]
2021	Acetowhite Lesion Segmentation	Colposcopic image	DL	Retrospective	Training: 9135 images; Validation: 609 images; Test: 609 images	No	AUC: 0.973; Accuracy: 0.906; Dice: 0.823	[55]
2021	CIN classification	H&E image	ML	Retrospective	Training: 100 patients; Validation: 5-fold CV; Test: 40 patients	No	AUC: 0.971; Accuracy: 90.0%	[186]
2019	Non-cancer vs cancer	DNA methylation; Gene expression	ML	Retrospective	Training: 420 samples; Validation: 10-fold CV; Test: 338 samples	No	Sensitivity: 96.2%; Specificity: 95.2%	[20]
	Prediction							

(*continued on next page*)

Table 1 (*continued*)

Year	Task	Modality	Model type	Study design	Size of dataset	Against Human	Performance	Ref.
2019	Pre-treatment prediction of response to NACT	MRI; Clinical data	ML	Retrospective	Total: 275 patients; Training: Test= 2:1	No	AUC: 0.999	[139]
2020	Actionable genetic alterations	H&E WSI	DL	Retrospective	Total: 261 patients; Validation: 3-fold CV	No	AUCs: *TCERG1:* 0.82; *STK11:* 0.81; *AMER1:* 0.76	[71]
2020	Subtype; Grade; Stage; Nodal status	MRI	ML	Retrospective	Total: 95 patients; Training: 80%; Validation: 20%	No	AUCs: Subtypes: 0.841; Tumor grades: 0.850; FIGO stages: 0.898; Nodal status: 0.879	[88]
2020	LNM	CT	ML	Retrospective	Total: 127 patients; Validation: leave-one-out CV	No	AUC: 0.841	[89]
2020	LNM	MRI; Clinical data	DL	Retrospective	Primary cohort: 338 patients; Validation: 141 patients	No	AUC: 0.933; Accuracy: 87.94%; Sensitivity: 90.62%; Specificity: 87.16%	[120]
2020	Parametrial invasion	MRI	ML	Retrospective	Total: 137 patients; Primary cohort: 95 patients; Validation: 10-fold CV	No	C-index: 0.941	[90]
2021	Determination of tumor functional uptake	PET	CNN	Retrospective	Training: 232 patients; Validation:5-fold CV	No	DSC: 0.80; Recall: 0.90	[187]
2021	Dose optimization during radiation therapy	2D/3D CT	DL	Retrospective	Training: 255 scans; Validation: 61 scans; Test: a.62 scans; b.30 scans	No	2D DSC: a. 0.88; b. 0.81; 3D DSC: a. 0.87; b. 0.82;	[188]
	Prognosis							
2019	PFS; OS	Clinical data	DL; Cox models	Retrospective	Total: 768 patients; Training: 8-fold; Validation: 1-fold; Test: 1-fold	No	C-indexes: PFS: 0.795; OS: 0.616	[189]
2019	Local relapse; Distant metastasis	^{18}F-FDG PET/CT	DL	Retrospective	Total: 142 patients	No	Local relapse: Accuracy: 89%; Sensitivity: 71%; Specificity: 93%; Distant metastasis: Accuracy: 87%; Sensitivity: 77%; Specificity: 90%	[151]
2021	DFI	^{18}F-FDG PET/CT	ML	Retrospective	Total:158 patients; Training: 80%; Validation: 5-fold CV; Test: 20%	No	AUC: 0.78	[152]
Uterine cancer								
	Screening							
2020	Screening in postmenopausal women	Serology: Metabolome	ML	Prospective	Training: 120 patients; Test: 1430 patients	No	Accuracy: 99.9%; Sensitivity: 100%; Specificity: 99.9%	[64]
	Diagnosis							
2019	Malignant vs Benign	MRI	ML	Retrospective	Total: 42 patients; Validation: leave-one-out CV	No	Accuracy: 91.7%; Sensitivity: 100%; Specificity: 90%	[72]
2019	Malignant vs Benign	MRI	ML	Retrospective	Training: 72 patients Validation: 10-fold CV	Yes	AUC: 0.93 Human: 0.68–0.80	[73]
2020	Malignant vs Benign;	H&E image	CNN	Retrospective	Total: 3302 images; Validation: 10-fold CV	Yes	AUC: 0.9579 Accuracy: 93.53% (Human: 88.33%); Sensitivity: 81.04% (Human: 89.48%); Specificity: 94.78% (Human: 88.06%)	[190]
2021	Malignant vs Benign	MRI	ML	Retrospective	Training: 104 patients; Test: 30 patients	Yes	AUC: 0.91 Human: 0.78–0.90	[74]
2021	Malignant vs Benign	US	ML	Retrospective	Total: 70 patients; Validation: 10-fold CV;	Yes	Accuracy: 85%; AUC: 0.86;	[75]
2021	Malignant/premalignant vs Benign;	Hysteroscopic image	CNN	Retrospective	Training: 6478 images; Validation: 250 images; Test: 250 images	Yes	Accuracy: 80.8%; Human: 64.4–72.8%	[94]
2021	Histologic & Molecular subtype; Gene mutations	H&E image	DCNN	Retrospective	a. 361 patients; b. 98 patients; c. 41 patients (A) a+b: Training: 80%; Validation: 10%;	No	AUCs: Histological: A. 0.969; c. 0.913; CNV-H: A. 0.934; c. 0.795; CNV-L: A. 0.889; c. 0.850;	[16]

(*continued on next page*)

Table 1 (*continued*)

Year	Task	Modality	Model type	Study design	Size of dataset	Against Human	Performance	Ref.
					Test: 10% (B) a for Training: 90%; Validation: 10%; b for Test Independent test: c		MSI-high: A. 0.827; c. 0.667; *TP53*: A. 0.873; c. 0.92	
2022	Malignant vs Benign	MRI	ML	Retrospective	Total: 172; Training: Test= 7:3	No	Accuracy: 87%; AUC: 0.96	[76]
	Prediction & Prognosis							
2017	DMI; Grade; LVSI	MRI	ML	Retrospective	Total: 137 patients; Validation: leave-one-out CV	Yes	AUCs: DMI: 0.84; LVSI: 0.80; High-grade: 0.83; DMI: Accuracy: 81.0% (Human: 83.2%); Sensitivity: 79.3% (Human: 84.5%); Specificity: 82.3% (Human: 82.3%);	[82]
2020	DMI; Lesion identification	MRI	DL	Retrospective	Training: 313 patients; Validation: 79 patients; Test: 138 patients	Yes	AUC: 0.78 Accuracy: 84.78% (Human: 78.3%); Sensitivity: 66.67% (Human: 61.1%); Specificity: 87.50% (Human: 80.8%);	[83]
2020	Grade; LVSI	MRI	ML	Retrospective	Total: 73 patients; Validation: leave-one-out CV	No	AUCs: High-grade: 0.64; LVSI: 0.59	[84]
2020	LNM	MRI	ML	Retrospective	Total: 622 patients; Validation: a. 351 patients; b. 271 patients	Yes	AUCs: Validation: a. 0.909; b. 0.885; Human: 0.623; 0.643;	[85]
2020	LNM; Prognosis: 5-year DSS	Clinical data	ML	Retrospective Prospective	Training: 763 patients; Test: a. 446 patients; b. 384 patients	No	AUCs: LNM: a. 0.82; 5-year DSS: a. 0.82; b. 0.84	[123]
2022	DMI; FIGO stage; Grade; LVSI	3D MRI	ML	Retrospective	Training: 94 patients; Test: 63 patients	Yes	AUCs: DMI: 0.81; FIGO stage: 0.84; High-grade: 0.74; LVSI: 0.80; Accuracies: DMI: 86% (Human: 79%); FIGO stage: 80% (Human: 78%)	[81]
2020	Chemotherapy benefit	Clinical data	ML	Retrospective	Training: 370 patients; Validation: 369 patients	No	C-index: Training: 0.67; Test: 0.73	[135]
2022	TME	WSI	Graph DL	Retrospective	Total: 548 patients; Validation: 5-fold CV	No	Accuracy: 0.727	[114]
Ovarian cancer								
	Screening							
2018	Non-cancer vs Benign vs Malignant	Serology: Postprandial increase in CA125	ML	-	Total: 551 patients; Validation: leave-one-out CV	No	Sensitivity: 91.7%; Specificity: 99.2%	[61]
2019	Screening	Serology: cfDNA	ML	-	Total: 243 patients; Training: 90%; Validation: 10-fold CV; Test: 10%	No	Sensitivity: 89%; Specificity: 98%	[66]
2022	Screening	Serology	ML	-	Training: 215 patients; Validation: 10-fold CV; Test: 54 patients	No	Sensitivity: 87%; Specificity: 98%	[62]
2022	Malignant vs Benign; Cancer stage	Serology: Methylation	ML	-	Training: 178 patients; Validation: 5-fold CV; Test: 184 patients	No	Accuracy: 89.5%	[22]
2022	Malignant vs Benign	Serology: Metabolome	DL	-	Training: 235 patients; Validation: 5-fold CV Test: 174 patients	No	AUC: 0.86	[63]
	Diagnosis							
2019	Malignant vs Benign; Subtype; Stage; Prognosis	Serology	ML	Retrospective	Training: 219 patients; Validation: 10-fold CV Test: 216 patients	No	Diagnosis: AUC: 0.968; Accuracy: 92.4%; Stage: AUC: 0.760; Accuracy: 69.0%	[97]

(*continued on next page*)

Table 1 (*continued*)

Year	Task	Modality	Model type	Study design	Size of dataset	Against Human	Performance	Ref.
2021	Malignant vs Benign;	US	DNN	Prospective	Training: 508 patients; Validation: 100 patients; Test: 150 patients	Yes	Sensitivity: 96%; Specificity:86.7%; (Human: 88.0%)	[191]
2022	Malignant vs Benign	US	DCNN	Retrospective	Training: 105532 patients; Validation: a. 868 patients; b. 335 patients; c. 889 patients	Yes	AUCs: a: 0.911; b. 0.870; c. 0.831; Accuracies: a. 88.8% (Human: 85.7%); b. 86.9% (Human: 81.1%)	[15]
2022	Malignant vs Benign	US	DL	Retrospective	Total: 422 patients; Training: 70%; Validation: 10%; Test: 20%	Yes	$DL_{feature}$: AUC: 0.93; Sensitivity: 92%; Specificity:85% $DL_{decision}$: AUC: 0.90; Sensitivity: 92%; Specificity:80% Human: AUC: 0.97 Sensitivity: 96%; Specificity:87%	[93]
2019	EOC vs Benign; Subtype of EOC; Prognosis: DFS	MRI	ML	Retrospective	Total: 286 patients; Leave-one-out CV: 195 patients; independent validation: 85 patients	Yes	Diagnosis: Accuracy: 90.6% (Human: 83.5%); Sensitivity: 90.3% (Human: 82.3%); Specificity: 91.3% (Human: 86.9%) Prognosis: AUC: 0.899 C-index: 0.755	[78]
2020	EOC vs Borderline	MRI	ML	Retrospective	Training: 250 patients; Validation: a.92 patients; b. 159 patients	Yes	AUCs: a. 0.932 (Human: 0.792); b. 0.902 (Human: 0.797)	[79]
2021	Subtype of EOC	MRI	ML	Retrospective	Training: 144 patients; Validation: a. 75 patients; b. 75 patients	No	AUCs: a. 0.806; b. 0.847	[77]
2021	Malignant vs Benign	MRI	CNN	Retrospective	Total: 451 patients; Training: 70%; Validation: 20%; Test: 10%	Yes	Accuracy: 87% (senior expert: 74%); Sensitivity:75% (senior expert: 83%); Specificity: 92% (senior expert: 70%)	[80]
2017	Subtype	WSI	ML	Retrospective	Training: 68 patients; Validation: leave-one-out CV; Test: 65 patients	Yes	Kappa: 0.89 (Human: 0.84–0.93)	[192]
2022	Subtype	H&E WSI	DL	Retrospective	Training: 485 patients; Validation: 3-fold CV; Test: 60 patients	No	C-index: 0.8097	[70]
2020	Grade; Prediction: Platinum response	WSI; Gene expression; Proteome	CNN	Retrospective	Total: 587 patients; Training: 80%; Validation: 5-fold CV; Test: 20%	No	AUCs: Diagnosis: > 0.95; Tumor grade: > 0.80; Chemotherapy response: 0.519–0.638	[96]
	Prediction							
2019	HRD for PARPi response; Platinum response; Prognosis: OS	WGS	ML	Retrospective	Total: 294 patients; Validation: CV	No	C-index: 0.64	[136]
2020	TME: T-cell exclusion	WSI; Transcriptome	ML	Retrospective	Training: 155 patients; Validation:84 patients; Test: 215 patients	No	-	[149]
2021	Combinatorial therapies	scRNA-seq; WGS; RNA-seq; Single-drug response profiling	ML	Retrospective	Total: 423 drug-targets; 110 point mutations; Training: 90% Validation: 10-fold CV Test: 10%	No	Accuracy: 91–94%	[150]
2022	Bevacizumab response	H&E WSI; Tissue microarray	DL	Retrospective	a. total: 288 patients; Training: 65%; Validation: 5-fold CV; Test: a. 35%; b. 175 patients	No	Accuracy: 0.882	[144]
2018	Complete cytoreduction at SCS; Prognosis: OS	Clinical data	ANN	Retrospective	Training: 155 patients; Validation: 39 patients	No	Importance of DFI: Complete cytoreduction: 0.231 OS: 0.306	[129]
2020	Complete cytoreduction	Clinical data	KNN	Retrospective	Training: 96 patients; Test: 51 patients	No	Accuracy: 66%	[130]

(*continued on next page*)

Table 1 (*continued*)

Year	Task	Modality	Model type	Study design	Size of dataset	Against Human	Performance	Ref.
2022	Risk stratification	WSI; CT; Clinical data; Gene expression	ML	Retrospective	Training: 404 patients; Validation: 4-fold CV; Test: 40 patients	No	C-index: 0.61	[142]
	Prognosis							
2018	Recurrence	Serology: CA125	ML	Retrospective	Total: 57 patients; Validation: 10-fold CV	No	Accuracy: 0.62	[193]
2018	3-year Recurrence	CT; Clinical data	DL	Retrospective	Training: 102 patients; Primary cohort: 49 patients; Validation: a. 49 patients; b. 45 patients	No	C-indexes: a. 0.713; b. 0.694 AUCs: a.0.772; b.0.825	[23]
2021	30-day Postoperative readmission & complication	CT scan reports; Clinical data	NLP; ML	Retrospective	Total: 291 patients; Validation: 5-fold CV	No	AUCs: Readmission: 0.70; Complication: 0.68	[155]
2021	3-year Recurrence	Gene expression	GA-XGBoost	Retrospective	a. 106 patients Training: Internal Validation= 2:1 External Validation: b. 73 patients; c. 50 patients; d. 23 patients	No	Sensitivity: 74–100%	[147]
Gynecologic oncology								
	Diagnosis							
2019	CUP	Gene expression	ML	Prospective	Training: 7791 patients; Validation: 5-fold CV; Test: 11644 patients	No	22 cancer types (EC; OC): Accuracy: 74.1%; Concordance: 67.4%	[100]
2019	Primary vs Metastatic	Gene expression	ML	Retrospective	Training: 10688 patients; Validation: 5-fold CV; Test: 211 patients	No	40 cancer types (CC; EC; OC; uterine carcinosarcoma): Accuracy: 97%	[101]
2020	CUP	Gene expression	CNN	Retrospective	Training: 18217 patients; Validation: 10-fold CV; Test: a. 394 patients; b. 69 patients; c. 23 patients	No	18 cancer types (OC uterine carcinosarcoma): Accuracy: 72.46%	[102]
2021	CUP	H&E WSI	DL	Retrospective	Training: 22833 patients; Validation: 6499 patients; Test: a. 682 patients; b. 743 patients	No	18 cancer types (CC; EC; OC): AUCs: Top-1: 0.80; Top-3: 0.93; Concordance: Top-1: 0.61; Top-3: 0.82	[99]
2022	CUP	WGS	ML	Retrospective	Training: 6082; Validation: 15-fold CV; Test: 674	No	35 cancer types (CC; EC; OC): Precision: ~90%	[103]
2022	CUP	Gene expression	DL; ML	Retrospective	Training: 8209 patients; Validation: 10-fold CV; Test: 3535 patients	No	33 cancer types (CC; EC; OC; uterine carcinosarcoma): AUC: 0.974 Concordance: Top-1: 0.73; Top-3: 0.90	[105]
	Prognosis							
2018	Chromatin organization ~ survival	Images of tumor cell nuclei	ML	Retrospective	Discovery cohort: 390 patients; Validation: 246 patients (OC); 354 patients (uterine sarcoma); 791 patients (EC)	No	-	[115]
2018	Shor-time mortality from chemotherapy start	EHR data	ML	Retrospective	Pan cancer: Derivation: 17832 patients; Validation: 9114 patients; Difference: 8718 patients	No	AUCs: 30-day mortality: CC: 0.955; OC: 0.956; 180-day mortality: CC: 0.888; OC: 0.895	[138]
2020	180-day mortality	EHR data	ML	Prospective	Gynecologic cancer: Test: 1049 patients	No	AUC: 0.91	[157]
2022	3-month mortality	EHR data	ML	Prospective	Genitourinary cancer: Test: 376 patients	Yes	AUC: 0.85; Sensitivity: 19% (Human: 20%)	[158]

Abbreviations: ANN, artificial neuronal network; AUC, area under curve; CA125, cancer antigen 125; CC, cervical cancer; cfDNA, cell-free DNA; C-index, concordance index; CIN, cervical intraepithelial neoplasia; CNN, convolutional network; CT, computed tomography; CUP, cancer of unknown primary; CV, cross-validation; DCNN, deep convolutional neural network; DFI, disease-free interval; DFS, disease-free survival; DL, deep learning; DMI, deep myometrial invasion; DNN, deep neural network; DSC, Dice similarity coefficient; DSS, disease-specific survival; EC, endometrial cancer; EHR, electronic health record; EOC, epithelial ovarian cancer; FIGO, the International Federation of Gynecology and Obstetrics; HPV, human papillomavirus; HRD, homologous recombination deficiency; H&E, hematoxylin and eosin-stained; KNN, k-nearest neighbor; LNM, lymph node metastasis; LVSI, lymphovascular space invasion; ML, machine learning; MRI, magnetic resonance imaging;

NACT, neoadjuvant chemotherapy; NC, not calculated; NLP, natural language processing; OC, ovarian cancer; OS, overall survival; PARPi, poly (ADP-ribose) polymerase inhibitor; PET, positron emission tomography; PFS, Progression-free survival; scRNA, single cell RNA; SCS, secondary cytoreductive surgery; SHG, second-harmonic generation; SIL, squamous epithelial lesion; TBS, the Bethesda system; THG, third-harmonic generation; TME, tumor microenvironment; US, ultrasonography; WGS, whole-genome sequencing; WSI, whole-slide images; ^{18}F-FDG PET/CT, ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography; -, Not mentioned or not available.

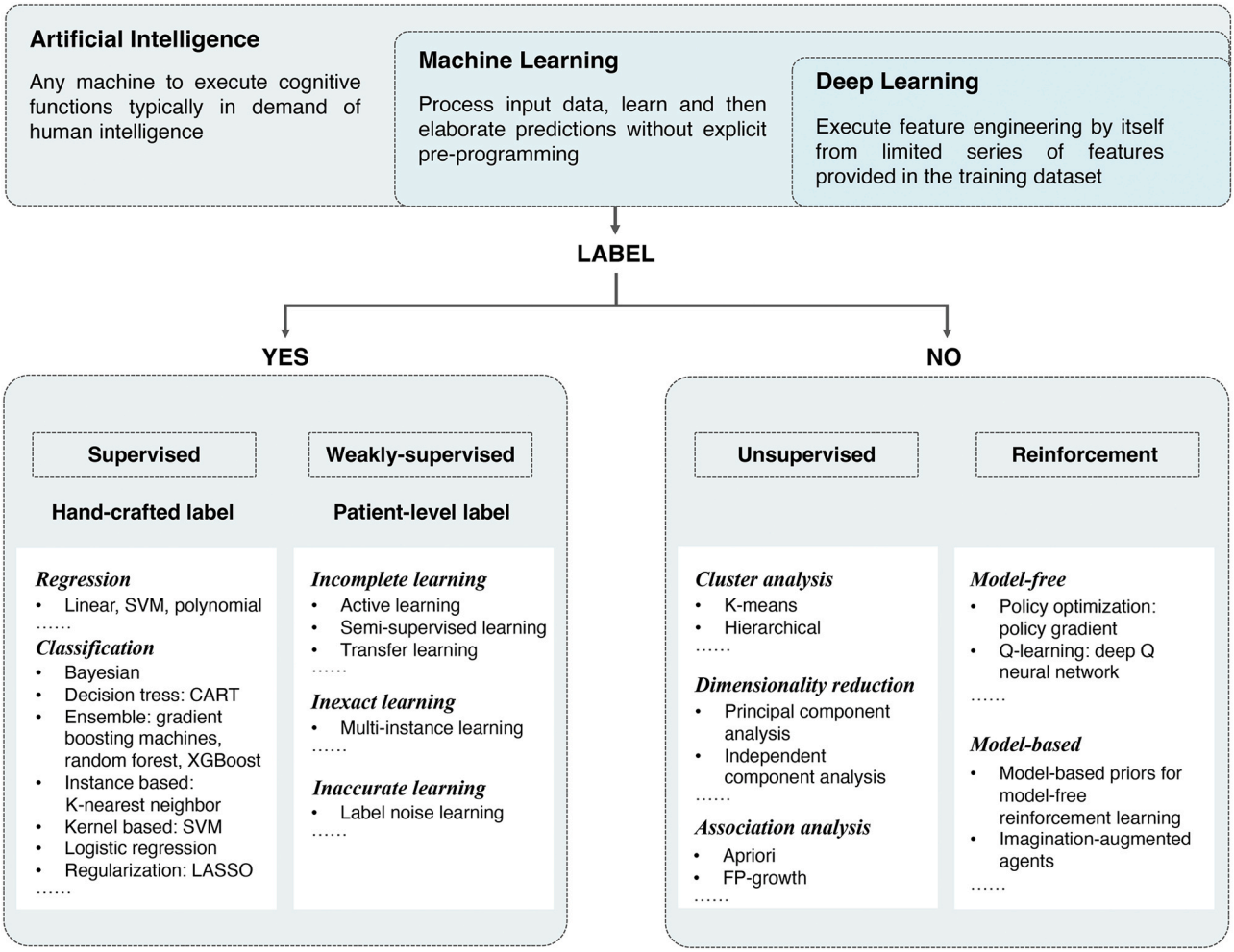


Fig. 1. Overview of artificial intelligence. AI is a vast term encompassing machine learning and deep learning. Data labeling is crucial for developing high-performance supervised learning models. Whenever source data is incomplete, inexact or inaccurate, weakly supervised models can be trained with patient-level annotations. Unsupervised approaches are capable of exploring unknown patterns in a collection of unlabeled data. Reinforcement learning is optimal to solve complex, sequential problems. CART, classification and regression trees; FP, frequent pattern; LASSO, least absolute shrinkage and selection operator; SVM, support vector machine.

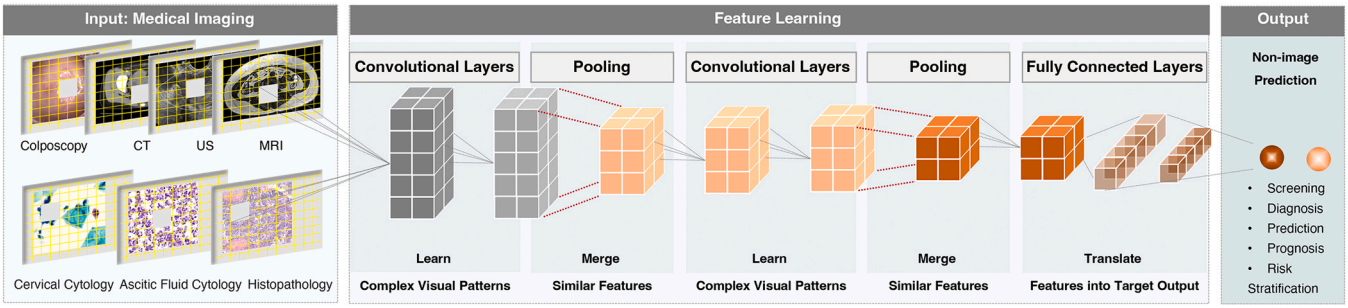
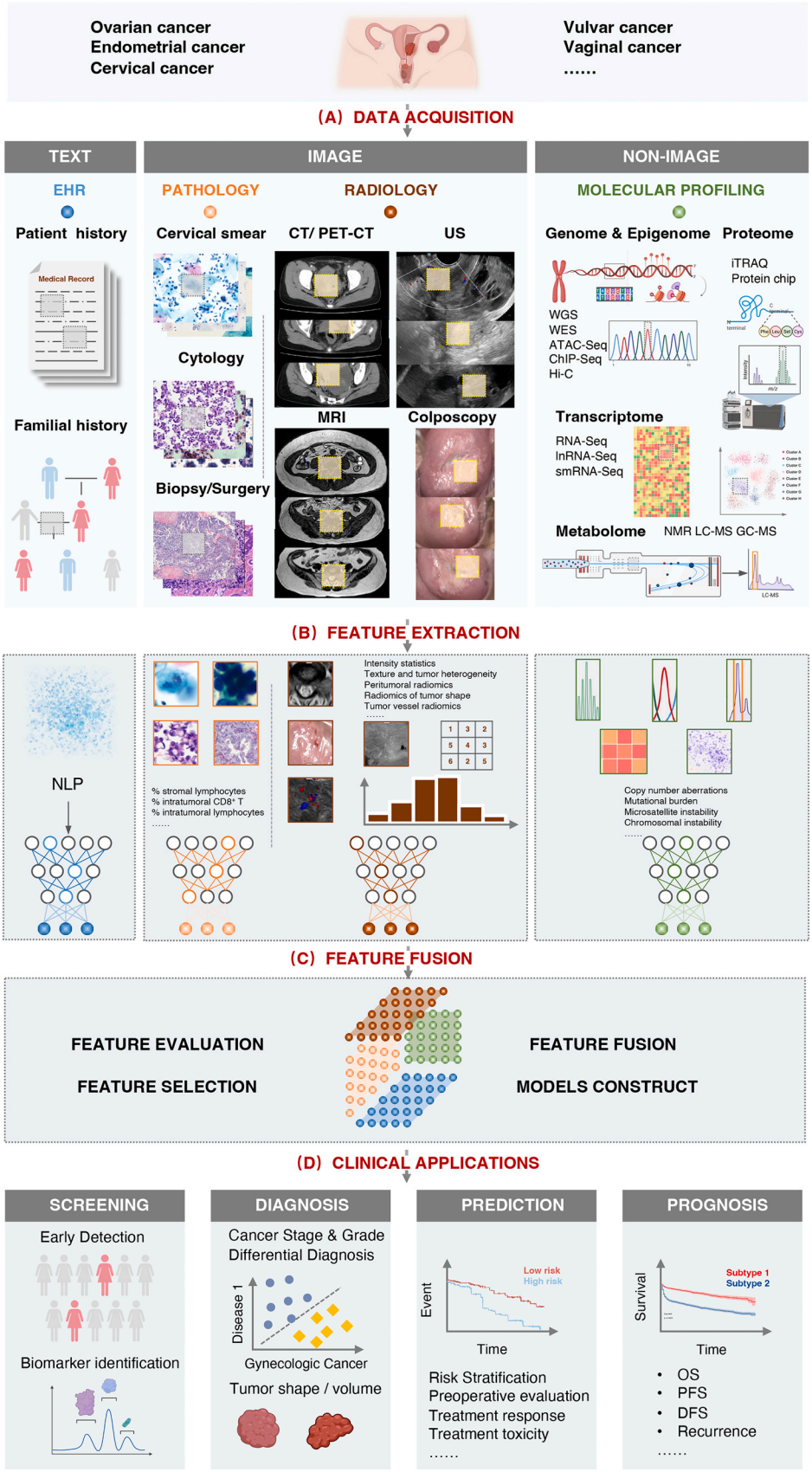


Fig. 2. General framework of convolutional neural network for medical imaging. Example of large-scale multilayer convolutional neural network takes various medical data as input, including colposcopic, radiologic, cytologic and histopathologic images. CNNs sequentially transform input images via convolutional, pooling and fully connected layers into flattened output vectors. CT, computed tomography; US, ultrasonography; MRI, magnetic resonance imaging.



(caption on next page)

Fig. 3. AI-enabled workflow for patients with gynecologic oncology. (A) Data acquisition. Patient data are processed by AI in the workflow across multiple modalities including text information, histopathology, radiology, and molecular profiling. (B) Feature extraction. AI learns a featurization in its lower-level towers for each modality. For EHR data, NLP is utilized to standardize data by mapping unstructured data from different hospitals to common format, and ultimately the entire record of a specific patient is generated using sequence models. (C) Feature fusion. The features extracted from each tower are merged into higher-level representations. (D) Clinical applications. The final outputs include screening, for example, quantitative analysis to speed up the identification of new biomarkers in support of large-scale patient screening, diagnosis (e.g., benign lesions versus malignant diseases), prediction (such as drug response), prognosis, among many others. DFS, disease-free survival; EHR, electronic health record; NLP, natural language processing; OS, overall survival; PFS, Progression-free survival. Created with BioRender.com.

into human genome is deemed as the major cause of cervical carcinogenesis, more and more cervical cancer prevention programs take advantage of HPV testing as a primary screening tool, followed with increased AI application in this aspect [42]. Recently, DeepHPV model was proposed to predict HPV integration genomic sites with an area under curve (AUC) of 0.634, and this model adding TCGA Pan Cancer attained a better AUC of 0.718 [43]. Real-time detection of the high-risk HPV 16 or 18 was realized by using DL-based microholography [44].

3.1.2. Cervical smear

On the flip side, inspired by the clinical criteria of Bethesda diagnostic system, a smear-level classifier was introduced with dual-path network for risk stratification in cervical cancer screening [45]. Based on massive whole-slide images (WSIs) of cytology, Cheng et al. developed a clinical-level-assisted diagnosis system verified by human–computer confrontation [46]. Additionally, the overall performance of an attention-guided convolutional network (CNN) was equivalent to that of pathologists with 10 years of practice experience [47]. Similarly, rapid TBS classification of the whole digital cervical liquid-based cytology smears (< 180 s/slides) was realized by a hybrid AI-steered diagnostic model at high specificity [48]. A cloud-based DLS was employed to provide point-of-care digital smear cytology for cervical cancer screening in resource-constrained areas [14]. In this proof-of-concept diagnostic study, compared to visual sample analysis, the DLS was feasible to detect squamous cell atypia especially high-grade atypia with high sensitivity. Moreover, it is a promising triage strategy to detect the co-expression of p16/ki67 before colposcopy referral [49]. Compared to conventional Pap cytology and manual dual stain (DS), CNN-augmented p16/ki67 DS could reduce colposcopy referrals by one-third [50].

3.1.3. Colposcopy

Moving to colposcopy, the main features and clues for colposcopist to grade CIN consist of acetowhite epithelium regions and appearance of blood vessel [51]. Several studies have investigated the usability and applicability of AI in the detection and classification of cervical lesions using colposcopic images [52–58]. An AI auxiliary diagnostic system exhibited satisfactory performance in grading colposcopic impressions and further directing multiple biopsy [52]. In other work, the comparison of machine learning algorithms against three colposcopists on different levels (intern, in-service and professional) reported the AI model exceeded the average human level and even obtained expert-level accuracy in CIN grading [53,54]. In fact, the turnaround time is long from biopsy to diagnostic report. Aimed to reduce time-to-result, a minimally invasive innovative approach was reported for visualizing cervical tissue three-dimensionally, employing nonlinear optics and near-infrared excitation [59]. Obviating the necessity of biopsy, fixation and staining, this ML-based method permitted accurate differentiation of invasive cancer, CIN and normal tissues with a weighted kappa coefficient of 0.86 [59]. Despite of few pitfalls, such technologies could chart a path for implementation of AI for gynecologic cancer screening particularly in resource-constrained settings.

3.2. Identification of novel screening biomarkers

Currently, a serum CA125 test is applied as the first-line screen for women at high risk of ovarian cancer [60]. But a ML-based study indicated CA125 postprandial increment exhibited diagnostic advantages over the conventional CA125 test [61]. In another promising study, a perception-based detection of high-grade serous ovarian cancer was constructed through machine learning from the spectral fingerprint of serum quantum-defect-modified nanotubes [62]. Compared with the present best clinical screening tests composed of CA125 and transvaginal ultrasound, the nanosensor-array technology improved the sensitivity to 87% with a specificity of 98% [62]. Particularly, the current known protein biomarkers were unable to achieve the predictive value of the proposed classifier, indicating the nanosensor array could be utilized as a discovery tool to quantify unidentified biomarkers.

Increasingly, analysis of data by AI from emerging minimally invasive techniques, such as liquid biopsies for metabolites, DNA from the Papanicolaou test, circulating tumor DNA (ctDNA) or cell-free DNA (cfDNA), offers considerable promise for detecting gynecologic cancers at an early stage when diseases can be effectively treated. Using a DL model, a seven-marker metabolite panel could be used to discriminate early-stage ovarian malignancy from benign pelvic diseases with an AUC of 0.86 [63]. Another work validated that using ML, a serum metabolomic-based screening could accurately detect endometrial cancer in postmenopausal women (accuracy, specificity, and sensitivity all > 99%) [64]. A multi-analyte blood test was created, named CancerSEEK, to screen eight types of cancer using combined asssys for circulating protein biomarkers and mutations in ctDNA [65]. Of note, using supervised ML, the sensitivity for detecting surgically resectable ovarian cancer reached 98% [65]. In general, cancer-derived cfDNA fragments are shorter and more variable than those found in healthy individuals. Cristiano and colleagues designed a machine learning model, which incorporated genome-wide cfDNA fragmentation features, to estimate the likelihood of a cfDNA mutation was related with 7 cancers (AUC 0.94) and specifically the model possessed a sensitivity 89% at 98% specificity in diagnosing ovarian cancer [66].

As the major epigenetic mark involved in carcinogenesis, DNA methylation has shown higher sensitivity relative to protein biomarkers and thereby may serve as a valuable source for developing novel biomarkers in the oncology care [67]. Aided by machine learning, the diagnostic power of several DNA methylation-based predictors was investigated, including four cervical cancer-specific DNA methylation markers (cg07211381, cg20708961, cg26490054 and cg12205729) with 95.2% specificity and 96.2% sensitivity [20], and blood-based differentially methylated regions for even predicting ovarian cancer stage [22]. Furthermore, the AI-enabled assay of plasma ctDNA methylation will be an alternative promising approach to improving risk stratification and informing better management of patients with gynecologic malignancies, which has been exemplified by the success in early diagnosis of lung cancer via ML-assisted ultrasensitive detection of ctDNA methylation patterns [68].

Altogether, as ever-growing data from genome, transcriptome, metabolome and proteome screening, artificial intelligence with high computing power may be capable of discovering new patterns and novel biomarkers for early gynecologic cancer detection, as well as reducing the healthcare costs and time required for diagnosis.

4. Making gynecologic cancer diagnosis, staging and grading more accurate

4.1. Cancer diagnosis, staging and grading

The key processes of diagnosis encompass the preclusion of gynecologic benign lesions and the further classification of gynecologic cancers in the aspect of histopathology, primary site and even molecular features. Cancer grading and staging, referring to estimate the tumor size, location and other factors, illuminate how serious and aggressive the malignancy is and ultimately are utilized to guide treatment recommendations. Applications of AI in the realm of gynecologic cancer diagnostics are discussed with focus on clinical utility, implications and future outlook (Fig. 3).

4.1.1. Cancer imaging

It has been extensively reported that AI-based models are deployed to improve the diagnostic accuracy and stratify tumor subtypes directly from cancer imaging involving histopathologic images, radiological images and so forth [69]. Clinically, the copy-number variation low (CNV-L) subtype of endometrial carcinoma carries the worst clinical outcomes in the molecular classification standard [18]. Given that the identification of histological subtypes, molecular subtyping and single-gene mutations is of paramount value for diagnosis and informing treatment strategies, a customized multi-resolution deep convolutional neural network (DCNN) was applied to predict the molecular subtypes of endometrial cancer as well as 18 common genes based on H&E-stained histopathological image without sequencing [16]. In ovarian cancer, a DL-based histotype determination was employed to automatically assess H&E-stained WSIs [70]. Interestingly, four cases from the external dataset, of which classification by AI were distinct from the reference diagnosis, were independently reviewed by two experts and the professional diagnoses were in accordance with the performance of AI [70]. Additionally, the application of AI trained on H&E histology slides could also predict standard histological biomarkers, detect pan-cancer genetic variants, further infer oncogenic drivers, molecular tumor subtypes and gene expression signatures for 14 of the most common cancer types [71]. This investigation indicated that *TCERG1*, *AMER1* and *STK11* alterations were detectable in patients with endocervical adenocarcinoma or cervical squamous cell carcinoma with high AUC values [71].

Radiological imaging is another mainstay of cancer diagnosis, staging and grading. Since uterine sarcomas are a rare type of uterine malignancies typically associated with poor prognosis, inadequate surgery in the event of occult sarcoma may deteriorate the clinical outcomes. As such, several effective AI-based diagnostic support tools are developed to differentiate uterine sarcomas from leiomyomas by combining radiomics and clinical parameters or ultrasonography (US) [72–76]. Superior performance is achieved over radiologists in diagnosis of uterine sarcomas [73,74]. Additionally, many AI-aided MRI radiomics models have been constructed to evaluate the nature of ovarian lesions and the discriminative performance, such as distinguishing between type I and type II epithelial ovarian cancer [77], approached or even surpassed radiologists [78–80]. Texture features extracted from MRI/CT imaging were also analyzed by AI to predict histological grading, FIGO staging, high-risk subtype, the presence of lymphovascular space invasion and deep myometrial invasion for endometrial cancer [81–87]; to differentiate clinicopathological signatures including lymph node status, parametrial invasion, histological subtypes, tumor grade and FIGO stage of cervical carcinoma [88–90].

In view of the safety, well tolerance and ubiquitous availability, US has been advocated as the most useful non-invasive tool in ovarian diseases identification [91]. Nonetheless, there are great variabilities observed in the subjective assessments of US images among radiologists, in spite of using established standard processes like the Ovarian-Adnexal Reporting and Data System (O-RADS) [92]. Two DL algorithms ($DL_{decision}$ and $DL_{feature}$) were modified from residual network in attempt to automatedly discern ovarian cancer from benign adnexal masses [93]. For ovarian malignancy, the diagnostic performance of $DL_{feature}$ (AUC 0.93) was on par with O-RADS (AUC 0.92, p 0.88) and expert assessment (AUC 0.97, p 0.07) [93]. In a multicentre diagnostic study, a DCNN model was developed in 105532 training dataset and tested in 2092 validation datasets containing one internal and two external validations [15]. In the internal and one of the external validation datasets, higher accuracies (88.8% and 86.9%, respectively) were achieved by the DCNN compared with average level of 16 (85.7%) and 19 radiologists (81.1%) respectively. Importantly, when DCNN-enable, the diagnostic accuracy of six junior radiologists was significantly improved from 78.3% to 87.6%, on par with the expert level.

Interestingly, a CNN-based model was constructed to automatically classify endometrial diseases from hysteroscopic images (accuracy: 80.8%), outperforming three gynecologists, and the overall performance of the CNN-assisted diagnosis by gynecologists was improved [94]. Thus, clinical auxiliary implementation of such AI models could not only help further training the next generation of experts by providing real-time guidance, but also ease the workload of skilled clinicians or alleviate the maldistribution of medical resources especially in developing countries.

4.1.2. Non-image data

Beyond cancer imaging data, as the molecular measurement techniques advance and the costs of these assays decline, data obtained from molecular profiling (such as transcriptome, metabolome, proteome and genome) are promising in diagnosis, staging and grading of gynecologic cancers. Using whole-exome sequencing data, a ML classifier was constructed to identify the heterogeneous, distinct composition of endometrioid ovarian carcinoma, which was a rare subcategory of epithelial ovarian tumors with incompletely disclosed molecular characteristics [95].

Furthermore, the multimodal analysis of imaging and non-imaging data, which possess complementary information, has the great potential to enable refined risk stratification and classify cancers into subtypes [19]. As an example, convolutional neural networks were developed to classify grade of serous ovarian carcinoma (AUC $>$ 0.80) through integrating proteomics, RNA-seq data and WSIs [96]. Unsupervised ML was proposed to identify the clinical stage of ovarian cancer patients based on preoperative blood biomarkers containing CA125, C-reactive protein (CRP), carbohydrate 19–9 (CA19–9), lactate dehydrogenase (LDH), albumin and et al. (AUC 0.76) and predict histotypes including high-grade serous, mucinous, clear cell and endometrioid and ones with AUCs of 0.785, 0.728, 0.650 and 0.597, respectively [97]. Moreover, the usage of ML revealed that among early-stage (stage I/II) EOC patients, two clusters were significantly correlated with worse prognosis [97].

4.2. Diagnosing cancers of unknown primary

In spite of extensive diagnostic investigation and clinical relevance, cancer of unknown primary (CUP) origin, accounting for 1–2% of all malignancies, poses enormous challenges owing to lack of tumor type-specific care alongside dramatic clinical deterioration [98]. Patient with CUP have to receive empirical therapies rather than effective site-specific treatment and are prone to quickly succumb to this aggressive disease, resulting in inferior prognosis [98]. To address these challenges, Tumor Origin Assessment via Deep Learning (TOAD) was adopted to concurrently determine the tumor as metastatic or primary and assign a primary differential based on routinely acquired whole-slide images corresponding to 18 cancer types [99]. For ovarian cancer and endometrial cancer, artificial intelligence was also applied to make accurate predictions of the cancer origin using targeted panel DNA sequencing data as input, with promising results [100]. Via prospective targeted sequencing assay of up to 468 key cancer-related gene, the molecular-oriented classifier was deployed to inform the site of origin for 95 of 141 patients with CUP across 22 tumor types [100].

Importantly, prospective application of the ML-based tool led to genomic-derived re-diagnosis of two patients with CUP and thereby prompted a change in clinical care plan, to which these patients ultimately responded [100].

Apart from genome-guided classification [101–103], it is well-established that tumor is partly an issue of developmental biology, indicating a new potential perspective from developmental origins [104]. After mapping 56 developmental trajectories of 33 cancer types (including ovarian, endometrial, cervical cancers and uterine carcinosarcoma), a Developmental Multilayer Perceptron (D-MLP) model was constructed through deconvoluting tumor transcriptomes into these component trajectories to reveal the cancer origin [105]. They further curated a cohort of 52 cases of CUP, which were refractory to standard pathology protocols and representative for most challenging cases. D-MLP classifier was capable of distinguishing such cases by developmental trajectories and clarify differential diagnoses [105]. Altogether, AI could yield accurate predictions, facilitate a better understanding of the carcinogenesis, and be utilized as an adjunct to or in lieu of auxiliary examination and extensive conventional workups to fast diagnose complicated cases of CUP, with the hope to improve the management and outcomes of CUP.

5. Prediction of patient prognosis and treatment recommendation

5.1. Risk stratification and prognosis

The choices of treatments consisting of surgery, chemotherapy, radiotherapy, targeted therapy and immunotherapy, vary tremendously among gynecologic tumors, depending on the clinical factors (such as age, stage, grade, and residual tumor volume after surgery), and functional omics indicators, for example, homologous recombination-deficient (HRD) for poly-ADP ribose polymerase (PARP) inhibitors in ovarian cancer [106,107], vascular endothelial growth factor (VEGF) for antiangiogenic therapy in ovarian and advanced cervical cancer [108,109], human epidermal growth factor receptor type 2 (HER2) for HER2-targeted therapy in ovarian and uterine cancer [110,111], microsatellite instability-high (MSI-H) and different mismatch repair (dMMR) for immune checkpoint therapy [112,113]. Therefore, in view of the enormous inter- and intra-tumoral heterogeneity in cancer risk, it is still a curial unmet need to determine patients with gynecologic cancers at risk of poor clinical outcomes. One notable example is that a graph DNN was leveraged to analyze the heterogeneous tumor environment on WSIs and determined a number of survival-related contexture features, which encompassed haemorrhagic necrosis, irregular growth pattern, intra-tumoral lymphocytic infiltration and so on, for endometrial, breast, lung and kidney cancers [114]. A similar study deployed ML algorithms to investigate the prognostic value of chromatin organization, termed Nucleotyping, in ~400000 images of tumor cell nuclei via image texture analysis, and found Nucleotyping was a significant predictor of outcomes in uterine sarcoma, ovarian and endometrial carcinomas [115].

Rapid developments of computational and experimental techniques provide opportunities to enrich complementary multi-omics data, and so AI will be well suitable to comprehensively dig into multiscale tumor information in support of precise prognostication and care. One group integrated five distinct data mining classification and ensemble learning to determine predictor variables for ovarian cancer patients, and identified FIGO staging, age, pathologic M and pathologic T as significant risk factors with regard to recurrence [116]. Overall, such risk prediction algorithms might allow physicians to provide precise risk stratification in order to improve prognostication for patients with gynecologic tumors.

5.2. Preoperative assessment

In the era of precision medicine, the application of AI, which resembles clinical reasoning, is well suited for preoperative risk stratification and prognostication in order to tailor personalized treatment planning and minimize exposure to unnecessary surgical intervention. Specifically, pelvic lymph node metastasis (LNM) constitutes one of the major prognostic factors for common cancers [117,118]. According to the International Federation of Gynecology and Obstetrics (FIGO) staging system, patients with cervical cancer who have either radiological or pathologic evidence of lymph node metastasis are diagnosed at late stage (IIIC), and these patients might benefit from chemoradiotherapy as the first option rather than surgery [119]. Therefore, a diagnostic study built an end-to-end DL model to estimate LNM probability in women with cervical cancer leveraging the high-resolution MRI, which was deemed as one of the most accurate imaging techniques [120]. In this research, a hybrid model combining clinicopathologic- and image-level information was able to preoperatively determine over 90% of cases with LNM with a specificity of 87.16%, and the output score from the hybrid model was correlated with the prognosis. Analogously, increasing enthusiasm and endeavors have been witnessed in preoperative evaluation of lymph-node status for the management of endometrial cancer patients, since two key prospective randomized clinical trials (RCTs) revealed that no benefits were obtained by routine systematic lymphadenectomy for early-stage endometrial cancer in terms of overall or disease-free survival while there was a substantial rise in incidence of long-term morbidity thenceforth [121,122]. Via incorporating molecular, clinical and histopathological biomarkers, a machine learning-based predictive model achieved high discriminative performance on predicting LNM of women with endometrial cancer in external validation cohort (AUC 0.82), i.e. identifying 55.8% of the individuals as $< 5\%$ risk for LNM [123].

Though chemotherapy remains the first-line treatment for recurrent ovarian cancer (ROC), increasing amount of evidence has indicated secondary cytoreductive surgery (SCS) may be of value to improving prognosis in selected population with platinum-sensitive [124–128], thus several work explored the ability of AI for complete cytoreduction prediction [129,130]. A clinical model using artificial neuronal network (ANN) based on 194 ROC patients to predict complete cytoreduction at the time of SCS, and found disease-free interval (importance: 0.231) was the most important factor predicting complete cytoreduction, followed by retroperitoneal recurrence (importance: 0.178), residual disease at primary surgical treatment (importance: 0.138) and FIGO staging (importance: 0.088) [129]. Furthermore, disease-free interval, complete cytoreduction and FIGO staging were significant predictors of overall survival (OS) for those patients (importance: 0.306, 0.217 and 0.100, respectively). Considering the prognostic significance of residual tumor status, an ovarian cancer-specific prediction model was proposed for pretreatment assessment of stages and histologic subtypes, and took advantage of an ordinal classification-based ML, extracting information from 32 clinical parameters available from preoperative blood tests, to segregate complete resection from optimal and suboptimal resection, which was not fully realized by standard classification algorithms [97].

5.3. Response to chemotherapy

Chemotherapy can improve prognosis in advanced-stage gynecologic cancer and lower the risk of relapse in early-stage disease, nevertheless, it remains challenging to balance benefits and considerable risks of chemotherapy [131–134]. For instance, a clinical calculator was developed using ML algorithms to predict the chemotherapy benefits for early-stage uterine cancer [135]. Likewise, using ML techniques, it was found that the homologous recombination deficiency was associated with better prognosis in ovarian cancer patients with platinum regimens [136]. Recent evidence showed that many cancer patients whose chemotherapies were often initiated too late in the disease trajectory

suffered from burdensome complications without potential benefits of chemotherapy and died soon after initiating treatment [137]. A EHR data-driven ML model was developed to estimate the short-term mortality, involving 30-day and 180-day mortality from the first day of new chemotherapy regimens, among ovarian and cervical cancer individuals administrating chemotherapy [138]. Model predictions were accurate for both of ovarian cancer (AUC 0.956 for 30-day mortality and 0.895 for 180-day mortality) and cervical cancer (AUC 0.955 for 30-day mortality and 0.888 for 180-day mortality). Importantly, the overall performance of the model was better than estimates from RCTs and the Surveillance, Epidemiology, and End Results (SEER) dataset, in which oncologists routinely seek for quantitative risk predictions [138]. Besides, MRI-based radiomic features harbored great potential to enable pretreatment prediction of the clinical response to neoadjuvant chemotherapy in patients with locally advanced cervical cancer, which permitted prudent selection of appropriate patients (AUC 0.999) [139].

Moreover, the high lethality of ovarian cancer is mainly owing to frequent recurrence along with increasingly resistance to chemotherapy [140]. Derived from gene expression signatures, biochemically-inspired ML was applied to predict the responses to cisplatin, carboplatin, and oxaliplatin for ovarian cancer patients [24]. For carboplatin, the best-performing ensemble of gene expression (designated Car1) was *AKT1, GNGT1, EIF3K, MTHFR, ERCC1, VEGFC, NEDD4L, VEGFB, NRAS, RAF1, TIGD1, SGK1, TP53, NLRP1,* and *GSR*, which predicted a long recurrence time in patients with ovarian cancer at an accuracy of 60% [24]. ML approaches integrated with bioinformatics method were also deployed to build a useful genetic interaction network and thereby to predict chemoresistance in ovarian cancer based on signature gene pairs with an AUC of 0.97 [141]. For HGSOC patients, a non-invasive recurrence prediction model, incorporating 16-dimensional DL feature and Cox proportional hazard regression, was developed to extract effective prognostic biomarkers from preoperative contrast enhanced CT images and could accurately predict 3-year recurrence in two validation datasets (AUC 0.772 and 0.825), which outperformed the accuracy attained by a clinical model with five known clinical characteristics (AUC 0.443 and 0.400) [23]. One notable example lies in that a ML model fusing clinicogenomic (HRD or homologous recombination proficiency status), radiologic (contrast-enhanced CT) and histopathological features demonstrated augmented predictive capacity for overall survival of HGSOC patients (concordance index: 0.61), providing a promising path towards better risk stratification of patients with gynecologic oncology [142].

5.4. Characterizing the tumor microenvironment and identifying candidates for other therapies

Bevacizumab, an anti-angiogenic monoclonal antibody, has been approved by the US FDA for maintenance treatment of recurrent ovarian cancer, nonetheless, only a small portion of patients at advanced stage could benefit from bevacizumab in combination with chemotherapy [143]. To address the issues mentioned above, an AI-based predictive model was present to accurately evaluate therapeutic effectiveness of bevacizumab according to digital histopathological images that were not pre-annotated by pathologists (accuracy: 0.882) [144]. Another clinically meaningful molecular marker for targeted therapy is HER2 aberration [145]. Via machine learning methods, expression-based HER2 index was utilized to precisely select appropriate patients who harbored HER2-enriched pattern and might gain the highest magnitude of benefits from HER2-targeted therapies [146]. Besides, based on gene expression profiles, ML models demonstrated that poor prognostic gene signature was closely related to protein kinase B (AKT) in ovarian cancer [147], indicating that targeting AKT might be a potentially effective anticancer strategy.

New therapeutic strategies and paradigms are in urgent need for ovarian cancer [17]. Revolutionized by the success of checkpoint inhibitors, the immune-based interventions have emerged as the powerful, most rapidly growing therapeutic pillars within oncology. Unfortunately, albeit approved in a variety of cancers by FDA, the immune checkpoint blockade therapy does not elicit existing and robust antitumor responses as expected in ovarian cancer [148]. The geospatial distribution of immune cells within tumor microenvironment, specifically, the close cancer–lymphocyte proximity, dramatically affects the efficacy of antitumor treatments such as immunotherapies. Thus, Desbois et al. deployed AI for integrated transcriptome and digital pathology analysis to molecularly characterize mediators of $CD8^+$ T cell exclusion in ovarian cancer [149]. This work uncovered two key hallmarks of immune-excluded phenotype, including high level of TGFβ and stromal activation and lack of tumor-derived antigen presentation, shedding light on TGFβ as a potential target. Moreover, a ML-based platform was implemented to tailor personalized cancer-selective combinatorial regimens for patients with HGSOC, which made use of functional single-drug response assay, whole-genome sequencing, bulk and single-cell RNA sequencing [150]. Overall, AI-based approaches will enable further insights and better understanding of the tumor microenvironment, facilitate the discovery of cancer-specific candidate drug targets, and thereby assist in identifying combination strategy, drug design and repurposing.

5.5. Follow up

Precise follow-up is one of the underexplored areas for AI application. Radiological assessment is an important part of follow-up care where ^{18}F-fluorodeoxyglucose positron emission tomography-computed tomography (^{18}F-FDG PET/CT) has been commonly employed for serially monitoring disease risk throughout treatment. The pre-treatment ^{18}F-FDG PET/CT radiomics in combination with AI were able to accurately predict local relapse, distant metastasis or disease-free survival for patients with locally advanced cervical cancer [151,152]. Given the current widespread adoption of comprehensive EHR in modern health systems, AI leveraging EHR data may complement clinician institution and timely triage cancer patients for escalated preventive care strategies or end-of-life care [153–158]. MEDomics, an oncology-oriented information technology infrastructure, enables continuous pan-cancer prognostication through integrating longitudinal comprehensive health data with real-world data [159]. The combined utilization of NLP with ML techniques was adopted for unstructured preoperative CT scan reports and then the predictive power was augmented for 30-day major postoperative complication (AUC 0.68) and hospital readmission (AUC 0.70) among ovarian cancer patients undergoing debulking surgery [155]. Furthermore, using an EHR-embedded ML algorithm [156], a prospective validation study was conducted to generate real-time, accurate predictions of 180-day mortality for 1049 outpatients with gynecologic cancers (AUC 0.91) [157]. Likewise, another prognostic study prospectively compared oncologist performance with a ML model trained with EHR data to predict 3-month mortality for outpatients (AUC 0.85) [158]. As a consequence, encounters flagged as being at high risk for short-term death were deemed appropriate for palliative or supportive measures like end-of-life conversations rather than aggressive treatment. Notably, both models surpass experts in prospective validation and their work stands for a remarkable early step in closing the gap between the health care ML and real-world cancer care.

6. Current challenges and future perspectives

Notwithstanding appreciable progress and tremendous success has been accomplished by oncological AI in the recent twenty years, some important challenges must be surmounted prior to the entry of AI to real-world scenarios.

6.1. Limitations

Most AI systems, particularly ML models are merely as good as their

background data. Broadly, the source data that are applied to train AI models carry massive inter-reader variability especially among non-experts facing rare and complex cases [160], inter- and intra-laboratory variability (such as differences in staining techniques, scanning characteristics for digital pathology, and process of sample preparation) [69], variability inherent to digital pictures (in consequence of distinct viewing angles, unevenness of exposure and so on) [7], and artificial variability introduced by data augmentation algorithms [161]. In addition, there are a wide range of data normalization (TPM, RPM, TMM, FPKM/RPKM) and data conversion (linear, log, etc.) utilized in multifarious programming languages [162,163]. These variabilities result in difficulties to define a standard reference against which the performance of models could be gauged [164,165].

Moving beyond imaging and omics data, an EHR of a major health care organization, which has the potential to gather the medical events of more than 10 million patients during a decade, is equivalent to roughly 200000 years of clinician wisdom and roughly 100 million years of patient prognostic data, spanning a fair amount of orphan diseases and disorders [166]. But the information-rich EHRs remain immensely underutilized, and the main reasons for this lie in unorganized structure and inconsistencies. One workable strategy is to standardize the codes and terminologies for specific disease in order to restructure individual data into easy-to-use data repository [167–169].

In parallel with the expeditious development of AI, ethical issues about the patient privacy and data security become growingly prominent in the era of big data. One of the critical strategies permitted by the rules governing the medical data privacy including the Health Insurance Portability and Accountability Act (HIPAA), is to deidentify data by means of concealing private information like names [170]. Unfortunately, poor preservation of privacy has precipitated due to the reidentification of individuals, facilitated by data triangulation from various sources, and the public-private collaborations for clinical translation [171,172]. Consequently, it is imperative to cautiously weigh between respect for patient privacy and benefits from big data-powered breakthroughs [173].

6.2. Potential biases

Large-scale datasets with clinical or manual annotations, especially expert annotations, were the essential basis for the development of fully supervised or semi-supervised artificial intelligence to a certain extent. Thus, the foremost limitation to constructing a supervised learning system for new medical tasks especially for imaging tasks is inadequate access to large and fully labeled datasets. Aside from being time consuming and laborious, hand-crafted methods might translate human bias to the AI models as well. Moreover, supervised learning-based models are restricted to features already understood or known by experts and the potentials to explore uncharted relevant features are impeded [25]. Thus, albeit easier to collect, small and labeled data for given tasks might lead to the poor generalizability of such AI models on unseen scenarios.

There are rising concerns over the risk that a trained AI system might magnify health disparities on account of biases lurking in the training dataset. It is worth to assess methodological quality of published work developed based upon AI because the drawbacks in the design, methods, conduct and result analysis might culminate in deviated evaluation of the model [174]. It is reported that only few deep learning studies were prospective trials, not to mention randomized trials [175]. Further validations are warranted in multicenter, prospective randomized trials to evaluate the actual performance of AI models.

6.3. Interpretability and uncertainty

The principal architecture of current AI models is designed with performance evaluated by accuracy, AUC, sensitivity, specificity and et al., at the expense of interpretability and transparency, resulting in AI is reputed as a black box [176]. From the clinical perspective, the main concerns lie in the physicians' trust in the AI-enabled indicators and the patients' confidence in the AI-supported medical care. While AI algorithms often outstrip the conventional interpretable models, the perceived deficiency in transparency and shortage of quantified uncertainty seriously undermine conviction in artificial intelligence, especially when encountering ambiguous clinical information. As such, endeavors to decode AI for clinicians and patients are warranted, and the emerging field of explainable AI (xAI), as evidenced by the surge in pertinent literature, will finally unlock the black box [177].

Indeed, some feasible tactics have been formulated from clinical and biological perspectives. Post-hoc interpretation methods, a prevalent category of xAI algorithms, are divided into local interpretation that valuate feature significance on a single input at a time, and global interpretation which seeks for the prominent characteristic combinations affecting the overall behaviors of AI models [177–180]. To yield patient-level interpretations, feature attribute strategies are attractive and practical by underlining salient feature importance. The techniques of feature attribute strategies include color mapping, which highlights areas of cancer images affecting the model outputs [181]; surrogate modelling techniques, which explicate the behavior of more sophisticated models by adopting simpler models [182]; and capturing image similarity that is commonly exhibited in the form of similar images frequently with superimposed gradient mapping [183]. To some extent, hand-crafted tools for cancer imaging could also furnish insights into the decision rules and actual logical flows of AI models. By way of illustration, vessel tortuosity metrics shed light on the biological and physical properties of tumor-associated neovasculature [184]. Besides, it is recommended to convey the degree of prediction uncertainty, as another dimension of transparency, which might arise from the collection and choice of the training data, reliability and integrality of data labeling, intrinsic biases of data sets, model misspecification and artifacts [185].

7. Conclusions

AI in gynecologic cancer care is an emerging research field. Looking ahead, the contributions of AI will span from recognition tasks of visually manifested conditions and clinical symptoms with expert-level or even preterhuman accuracy, to more innovative utilization like cancer prognostication and eventually personalized therapy and drug discovery. Despite the indisputable potential of oncologic AI, the major ethical and technical problems involving the integration of multimodal data and the deciphering the black box of AI become increasingly prominent. The true clinical value of medical AI will not be achieved in real-world cancer care until the pivotal issues are systematically resolved.

CRediT authorship contribution statement

Literature review and data collection were performed by Yuting Jiang and Chengdi Wang. The first draft of the review was written by Yuting Jiang, Chengdi Wang, and Shengtao Zhou. Yuting Jiang, Chengdi Wang, and Shengtao Zhou participated in editing tables and figures. All authors approved the final manuscript. Data availability No data was used for the research described in the article.

Declaration of Competing Interest

We declare no conflict of interest.

Data availability

No data was used for the research described in the article.

Acknowledgements

This work was supported by grants from National Natural Science Foundation of China (grant #82273255, #81822034, #81821002, and #81773119 to S.Z), National Key Research and Development Program of China (grant #2022YFA1106600, #2017YFA0106800 and #2018YFA0109200 to S.Z), Sichuan Science-Technology Project (grant #22ZYZYTS0070 and #2019YFH0144 to S.Z), and Direct Scientific Research Grants from West China Second Hospital, Sichuan University (grant #KS021 and #K1907 to S.Z).

References

[1] B. Meskó, M. Görög, A short guide for medical professionals in the era of artificial intelligence, npj Digit. Med. 3 (1) (2020) 1–8.
[2] P. Wang, On defining artificial intelligence, J. Artif. Gen. Intell. 10 (2) (2019) 1–37.
[3] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[4] W. Bulten, H. Pinckaers, H. van Boven, R. Vink, T. de Bel, B. van Ginneken, et al., Automated deep-learning system for Gleason grading of prostate cancer using biopsies: a diagnostic study, Lancet Oncol. 21 (2) (2020) 233–241.
[5] W. Bulten, K. Kartasalo, P.C. Chen, P. Ström, H. Pinckaers, K. Nagpal, et al., Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge, Nat. Med. 28 (1) (2022) 154–163.
[6] Y. Liu, A. Jain, C. Eng, D.H. Way, K. Lee, P. Bui, et al., A deep learning system for differential diagnosis of skin diseases, Nat. Med. 26 (6) (2020) 900–908.
[7] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, et al., Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (7639) (2017) 115–118.
[8] H.A. Haenssle, C. Fink, F. Toberer, J. Winkler, W. Stolz, T. Deinlein, et al., Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions, Ann. Oncol. 31 (1) (2020) 137–143.
[9] B.H. Kann, A. Hosny, H. Aerts, Artificial intelligence for clinical oncology, Cancer Cell 39 (7) (2021) 916–927.
[10] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, et al., Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA Cancer J. Clin. 71 (3) (2021) 209–249.
[11] W.H.O, Global strategy to accelerate the elimination of cervical cancer as a public health problem, 2020. ⟨https://www.who.int/publications/i/item/9789240014107⟩.
[12] P.A. Cohen, A. Jhingran, A. Oaknin, L. Denny, Cervical cancer, Lancet 393 (10167) (2019) 169–182.
[13] K. Alfaro, M. Maza, M. Cremer, R. Masch, M. Soler, Removing global barriers to cervical cancer prevention and moving towards elimination, Nat. Rev. Cancer 21 (10) (2021) 607–608.
[14] O. Holmström, N. Linder, H. Kaingu, N. Mbuuko, J. Mbete, F. Kinyua, et al., Point-of-care digital cytology with artificial intelligence for cervical cancer screening in a resource-limited setting, JAMA Netw. Open 4 (3) (2021), e211740.
[15] Y. Gao, S. Zeng, X. Xu, H. Li, S. Yao, K. Song, et al., Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in China: a retrospective, multicentre, diagnostic study, Lancet Digit. Health 4 (3) (2022) e179–e187.
[16] R. Hong, W. Liu, D. DeLair, N. Razavian, D. Fenyö, Predicting endometrial cancer subtypes and molecular features from histopathology images using multi-resolution deep learning models, Cell Rep. Med. 2 (9) (2021), 100400.
[17] S. Lheureux, M. Braunstein, A.M. Oza, Epithelial ovarian cancer: evolution of management in the era of precision medicine, CA Cancer J. Clin. 69 (4) (2019) 280–304.
[18] The Cancer Genome Atlas Research Network, Integrated genomic characterization of endometrial carcinoma, Nature 497 (7447) (2013) 67–73.
[19] K.M. Boehm, P. Khosravi, R. Vanguri, J. Gao, S.P. Shah, Harnessing multimodal data integration to advance precision oncology, Nat. Rev. Cancer 22 (2) (2022) 114–126.
[20] W. Xu, M. Xu, L. Wang, W. Zhou, R. Xiang, Y. Shi, et al., Integrative analysis of DNA methylation and gene expression identified cervical cancer-specific diagnostic biomarkers, Signal. Transduct. Target. Ther. 4 (1) (2019) 55.
[21] K. Bera, K.A. Schalper, D.L. Rimm, V. Velcheti, A. Madabhushi, Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology, Nat. Rev. Clin. Oncol. 16 (11) (2019) 703–715.
[22] N. Li, X. Zhu, W. Nian, Y. Li, Y. Sun, G. Yuan, et al., Blood-based DNA methylation profiling for the detection of ovarian cancer, Gynecol. Oncol. 167 (2) (2022) 295–305.
[23] S. Wang, Z. Liu, Y. Rong, B. Zhou, Y. Bai, W. Wei, et al., Deep learning provides a new computed tomography-based prognostic biomarker for recurrence prediction in high-grade serous ovarian cancer, Radiother. Oncol. 132 (2019) 171–177.
[24] E.J. Mucaki, J.Z.L. Zhao, D.J. Lizotte, P.K. Rogan, Predicting responses to platin chemotherapy agents with biochemically-inspired machine learning, Signal. Transduct. Target. Ther. 4 (1) (2019) 1.
[25] J. Lipkova, R.J. Chen, B. Chen, M.Y. Lu, M. Barbieri, D. Shao, et al., Artificial intelligence for multimodal data integration in oncology, Cancer Cell 40 (10) (2022) 1095–1110.
[26] B. Bhinder, C. Gilvary, N.S. Madhukar, O. Elemento, Artificial intelligence in cancer research and precision medicine, Cancer Discov. 11 (4) (2021) 900–915.
[27] G. Cammarota, G. Ianiro, A. Ahern, C. Carbone, A. Temko, M.J. Claesson, et al., Gut microbiome, big data and machine learning to promote precision medicine for cancer, Nat. Rev. Gastroenterol. Hepatol. 17 (10) (2020) 635–648.
[28] N.D. Shah, E.W. Steyerberg, D.M. Kent, Big data and predictive analytics: recalibrating expectations, JAMA 320 (1) (2018) 27–28.
[29] C. Gong, J. Yang, J. You, M. Sugiyama, Centroid estimation with guaranteed efficiency: a general framework for weakly supervised learning, IEEE Trans. Pattern Anal. Mach. Intell. 44 (6) (2022) 2841–2855.
[30] M. Alloghani, D. Al-Jumeily, J. Mustafina, A. Hussain, A.J. Aljaaf, A systematic review on supervised and unsupervised machine learning algorithms for data science, Supervised unsupervised Learn. data Sci. (2020) 3–21.
[31] M. Botvinick, S. Ritter, J.X. Wang, Z. Kurth-Nelson, C. Blundell, D. Hassabis, Reinforcement learning, fast and slow, Trends Cogn. Sci. 23 (5) (2019) 408–422.
[32] S. Nath, E. Korot, D.J. Fu, G. Zhang, K. Mishra, A.Y. Lee, et al., Reinforcement learning in ophthalmology: potential applications and challenges to implementation, Lancet Digit. Health 4 (9) (2022) e692–e697.
[33] X. Gao, J. Si, Y. Wen, M. Li, H. Huang, Reinforcement learning control of robotic knee with human-in-the-loop by flexible policy iteration, IEEE Trans. Neural Netw. Learn. Syst. 33 (10) (2022) 5873–5887.
[34] O.Y. Kudsi, F. Gokcal, N. Bou-Ayash, A.S. Crawford, K. Chang, A. Chudner, et al., Robotic ventral hernia repair: lessons learned from a 7-year experience, Ann. Surg. 275 (1) (2022) 9–16.
[35] M. Wainberg, D. Merico, A. Delong, B.J. Frey, Deep learning in biomedicine, Nat. Biotechnol. 36 (9) (2018) 829–838.
[36] K.L. Kehl, H. Elmarakeby, M. Nishino, E.M. Van Allen, E.M. Lepisto, M.J. Hassett, et al., Assessment of deep natural language processing in ascertaining oncologic outcomes from radiology reports, JAMA Oncol. 5 (10) (2019) 1421–1429.
[37] J.T. Henderson, E.M. Webber, G.F. Sawaya, Screening for ovarian cancer: updated evidence report and systematic review for the us preventive services task force, JAMA 319 (6) (2018) 595–606.
[38] J. Melnikow, J.T. Henderson, B.U. Burda, C.A. Senger, S. Durbin, M.S. Weyrich, Screening for cervical cancer with high-risk human papillomavirus testing: updated evidence report and systematic review for the us preventive services task force, JAMA 320 (7) (2018) 687–705.
[39] E.J. Crosbie, S.J. Kitson, J.N. McAlpine, A. Mukhopadhyay, M.E. Powell, N. Singh, Endometrial cancer, Lancet 399 (10333) (2022) 1412–1428.
[40] R.B.S. Roden, P.L. Stern, Opportunities and challenges for human papillomavirus vaccination in cancer, Nat. Rev. Cancer 18 (4) (2018) 240–254.
[41] E.T.H. Fontham, A.M.D. Wolf, T.R. Church, R. Etzioni, C.R. Flowers, A. Herzig, et al., Cervical cancer screening for individuals at average risk: 2020 guideline update from the American Cancer Society, CA Cancer J. Clin. 70 (5) (2020) 321–346.
[42] C. Marth, F. Landoni, S. Mahner, M. McCormack, A. Gonzalez-Martin, N. Colombo, Cervical cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up, Ann. Oncol. 28 (suppl_4) (2017) iv72–iv83.
[43] R. Tian, P. Zhou, M. Li, J. Tan, Z. Cui, W. Xu, et al., DeepHPV: a deep learning model to predict human papillomavirus integration sites, Brief. Bioinform 22 (4) (2021).
[44] D. Pathania, C. Landeros, L. Rohrer, V. D'Agostino, S. Hong, I. Degani, et al., Point-of-care cervical cancer screening using deep learning-based microholography, Theranostics 9 (26) (2019) 8438–8447.
[45] H. Lin, H. Chen, X. Wang, Q. Wang, L. Wang, P.A. Heng, Dual-path network with synergistic grouping loss and evidence driven risk stratification for whole slide cervical image analysis, Med. Image Anal. 69 (2021), 101955.
[46] S. Cheng, S. Liu, J. Yu, G. Rao, Y. Xiao, W. Han, et al., Robust whole slide image analysis for cervical cancer screening using deep learning, Nat. Commun. 12 (1) (2021) 5639.
[47] L. Cao, J. Yang, Z. Rong, L. Li, B. Xia, C. You, et al., A novel attention-guided convolutional network for the detection of abnormal cervical cells in cervical cancer screening, Med. Image Anal. 73 (2021), 102197.
[48] X. Zhu, X. Li, K. Ong, W. Zhang, W. Li, L. Li, et al., Hybrid AI-assistive diagnostic model permits rapid TBS classification of cervical liquid-based thin-layer cell smears, Nat. Commun. 12 (1) (2021) 3541.
[49] P. Giorgi Rossi, F. Carozzi, G. Ronco, E. Allia, S. Bisanzi, A. Gillio-Tos, et al., P16/ki67 and E6/E7 mRNA accuracy and prognostic value in triaging HPV DNA-positive women, J. Natl. Cancer Inst. 113 (3) (2021) 292–300.
[50] N. Wentzensen, B. Lahrmann, M.A. Clarke, W. Kinney, D. Tokugawa, N. Poitras, et al., Accuracy and efficiency of deep-learning-based automation of dual stain cytology in cervical cancer screening, J. Natl. Cancer Inst. 113 (1) (2021) 72–79.
[51] H. Sahlgren, K.M. Elfström, H. Lamin, A. Carlsten-Thor, C. Eklund, J. Dillner, et al., Colposcopic and histopathologic evaluation of women with HPV persistence exiting an organized screening program, Am. J. Obstet. Gynecol. 222 (3) (2020), 253.e1-e8.
[52] P. Xue, C. Tang, Q. Li, Y. Li, Y. Shen, Y. Zhao, et al., Development and validation of an artificial intelligence system for grading colposcopic impressions and guiding biopsies, BMC Med 18 (1) (2020) 406.
[53] Y. Li, J. Chen, P. Xue, C. Tang, J. Chang, C. Chu, et al., Computer-aided cervical cancer diagnosis using time-lapsed colposcopic images, IEEE Trans. Med. Imaging 39 (11) (2020) 3403–3415.

[54] Y. Li, Z.H. Liu, P. Xue, J. Chen, K. Ma, T. Qian, et al., GRAND: a large-scale dataset and benchmark for cervical intraepithelial neoplasia grading with fine-grained lesion description, Med. Image Anal. 70 (2021), 102006.
[55] Z. Yue, S. Ding, X. Li, S. Yang, Y. Zhang, Automatic acetowhite lesion segmentation via specular reflection removal and deep attention network, IEEE J. Biomed. Health Inf. 25 (9) (2021) 3529–3540.
[56] P. Xue, M.T.A. Ng, Y. Qiao, The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence, BMC Med 18 (1) (2020) 169.
[57] L. Hu, D. Bell, S. Antani, Z. Xue, K. Yu, M.P. Horning, et al., An observational study of deep learning and automated evaluation of cervical images for cancer screening, J. Natl. Cancer Inst. 111 (9) (2019) 923–932.
[58] D. Song, E. Kim, X. Huang, J. Patruno, H. Munoz-Avila, J. Heflin, et al., Multimodal entity coreference for cervical dysplasia diagnosis, IEEE Trans. Med. Imaging 34 (1) (2015) 229–245.
[59] T. Matsui, R. Tamoto, A. Iwasa, M. Mimura, S. Taniguchi, T. Hasegawa, et al., Nonlinear optics with near-infrared excitation enable real-time quantitative diagnosis of human cervical cancers, Cancer Res 80 (17) (2020) 3745–3754.
[60] O. Blyuss, M. Burnell, A. Ryan, A. Gentry-Maharaj, I.P. Mariño, J. Kalsi, et al., Comparison of longitudinal CA125 algorithms as a first-line screen for ovarian cancer in the general population, Clin. Cancer Res 24 (19) (2018) 4726–4733.
[61] Z. Gu, Y. He, Y. Zhang, M. Chen, K. Song, Y. Huang, et al., Postprandial increase in serum CA125 as a surrogate biomarker for early diagnosis of ovarian cancer, J. Transl. Med 16 (1) (2018) 114.
[62] M. Kim, C. Chen, P. Wang, J.J. Mulvey, Y. Yang, C. Wun, et al., Detection of ovarian cancer via the spectral fingerprinting of quantum-defect-modified carbon nanotubes in serum by machine learning, Nat. Biomed. Eng. 6 (3) (2022) 267–275.
[63] E. Irajizad, C.Y. Han, J. Celestino, R. Wu, E. Murage, R. Spencer, et al., A blood-based metabolite panel for distinguishing ovarian cancer from benign pelvic masses, Clin. Cancer Res. 28 (21) (2022) 4669–4676.
[64] J. Troisi, A. Raffone, A. Travaglino, G. Belli, C. Belli, S. Anand, et al., Development and validation of a serum metabolomic signature for endometrial cancer screening in postmenopausal women, JAMA Netw. Open 3 (9) (2020), e2018327.
[65] J.D. Cohen, L. Li, Y. Wang, C. Thoburn, B. Afsari, L. Danilova, et al., Detection and localization of surgically resectable cancers with a multi-analyte blood test, Science 359 (6378) (2018) 926–930.
[66] S. Cristiano, A. Leal, J. Phallen, J. Fiksel, V. Adleff, D.C. Bruhm, et al., Genome-wide cell-free DNA fragmentation in patients with cancer, Nature 570 (7761) (2019) 385–389.
[67] M.L. Cheng, E. Pectasides, G.J. Hanna, H.A. Parsons, A.D. Choudhury, G. R. Oxnard, Circulating tumor DNA in advanced solid tumors: clinical relevance and future directions, CA Cancer J. Clin. 71 (2) (2021) 176–190.
[68] N. Liang, B. Li, Z. Jia, C. Wang, P. Wu, T. Zheng, et al., Ultrasensitive detection of circulating tumour DNA via deep methylation sequencing aided by machine learning, Nat. Biomed. Eng. 5 (6) (2021) 586–599.
[69] J. van der Laak, G. Litjens, F. Ciompi, Deep learning in histopathology: the path to the clinic, Nat. Med. 27 (5) (2021) 775–784.
[70] H. Farahani, J. Boschman, D. Farnell, A. Darbandsari, A. Zhang, P. Ahmadvand, et al., Deep learning-based histotype diagnosis of ovarian carcinoma whole-slide pathology images, Mod. Pathol. 35 (12) (2022) 1983–1990.
[71] J.N. Kather, L.R. Heij, H.I. Grabsch, C. Loeffler, A. Echle, H.S. Muti, et al., Pan-cancer image-based detection of clinically actionable genetic alterations, Nat. Cancer 1 (8) (2020) 789–799.
[72] M. Malek, M. Gity, A. Alidoosti, Z. Oghabian, P. Rahimifar, S.M. Seyed Ebrahimi, et al., A machine learning approach for distinguishing uterine sarcoma from leiomyomas based on perfusion weighted MRI parameters, Eur. J. Radiol. 110 (2019) 203–211.
[73] M. Nakagawa, T. Nakaura, T. Namimoto, Y. Iyama, M. Kidoh, K. Hirata, et al., Machine Learning to differentiate T2-weighted hyperintense uterine leiomyomas from uterine sarcomas by utilizing multiparametric magnetic resonance quantitative imaging features, Acad. Radiol. 26 (10) (2019) 1390–1399.
[74] T. Wang, J. Gong, Q. Li, C. Chu, W. Shen, W. Peng, et al., A combined radiomics and clinical variables model for prediction of malignancy in T2 hyperintense uterine mesenchymal tumors on MRI, Eur. Radiol. 31 (8) (2021) 6125–6135.
[75] V. Chiappa, M. Interlenghi, C. Salvatore, F. Bertolina, G. Bogani, A. Ditto, 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. 161 (3) (2021) 838–844.
[76] M. Dai, Y. Liu, Y. Hu, G. Li, J. Zhang, Z. Xiao, et al., Combining multiparametric MRI features-based transfer learning and clinical parameters: application of machine learning for the differentiation of uterine sarcomas from atypical leiomyomas, Eur. Radiol. 32 (11) (2022) 7988–7997.
[77] J. Jian, Y. Li, P.J. Pickhardt, W. Xia, Z. He, R. Zhang, et al., MR image-based radiomics to differentiate type I and type II epithelial ovarian cancers, Eur. Radiol. 31 (1) (2021) 403–410.
[78] H. Zhang, Y. Mao, X. Chen, G. Wu, X. Liu, P. Zhang, et al., Magnetic resonance imaging radiomics in categorizing ovarian masses and predicting clinical outcome: a preliminary study, Eur. Radiol. 29 (7) (2019) 3358–3371.
[79] Y. Li, J. Jian, P.J. Pickhardt, F. Ma, W. Xia, H. Li, et al., MRI-based machine learning for differentiating borderline from malignant epithelial ovarian tumors: a multicenter study, J. Magn. Reson. Imaging 52 (3) (2020) 897–904.
[80] R. Wang, Y. Cai, I.K. Lee, R. Hu, S. Purkayastha, I. Pan, et al., Evaluation of a convolutional neural network for ovarian tumor differentiation based on magnetic resonance imaging, Eur. Radiol. 31 (7) (2021) 4960–4971.
[81] T.L. Lefebvre, Y. Ueno, A. Dohan, A. Chatterjee, M. Vallières, E. Winter-Reinhold, et al., Development and validation of multiparametric MRI-based radiomics models for preoperative risk stratification of endometrial cancer, Radiology 305 (2) (2022), 212873.
[82] Y. Ueno, B. Forghani, R. Forghani, A. Dohan, X.Z. Zeng, F. Chamming's, et al., Endometrial carcinoma: MR imaging-based texture model for preoperative risk stratification-a preliminary analysis, Radiology 284 (3) (2017) 748–757.
[83] X. Chen, Y. Wang, M. Shen, B. Yang, Q. Zhou, Y. Yi, et al., Deep learning for the determination of myometrial invasion depth and automatic lesion identification in endometrial cancer MR imaging: a preliminary study in a single institution, Eur. Radiol. 30 (9) (2020) 4985–4994.
[84] M. Bereby-Kahane, R. Dautry, E. Matzner-Lober, F. Cornelis, D. Sebbag-Sfez, V. Place, et al., Prediction of tumor grade and lymphovascular space invasion in endometrial adenocarcinoma with MR imaging-based radiomic analysis, Diagn. Interv. Imaging 101 (6) (2020) 401–411.
[85] B.C. Yan, Y. Li, F.H. Ma, G.F. Zhang, F. Feng, M.H. Sun, et al., Radiologists with MRI-based radiomics aids to predict the pelvic lymph node metastasis in endometrial cancer: a multicenter study, Eur. Radiol. 31 (1) (2021) 411–422.
[86] S. Ytre-Hauge, J.A. Dybvik, A. Lundervold, O. Salvesen Ø, C. Krakstad, K. E. Fasmer, et al., Preoperative tumor texture analysis on MRI predicts high-risk disease and reduced survival in endometrial cancer, J. Magn. Reson. Imaging 48 (6) (2018) 1637–1647.
[87] W. Yue, N. Meng, J. Wang, W. Liu, X. Wang, M. Yan, et al., Comparative analysis of the value of diffusion kurtosis imaging and diffusion-weighted imaging in evaluating the histological features of endometrial cancer, Cancer Imaging 19 (1) (2019) 9.
[88] M. Wang, J.A.U. Perucho, K.Y. Tse, M.M.Y. Chu, P. Ip, E.Y.P. Lee, MRI texture features differentiate clinicopathological characteristics of cervical carcinoma, Eur. Radiol. 30 (10) (2020) 5384–5391.
[89] X. Chen, W. Liu, T.C. Thai, T. Castellano, C.C. Gunderson, K. Moore, et al., Developing a new radiomics-based CT image marker to detect lymph node metastasis among cervical cancer patients, Comput. Methods Prog. Biomed. 197 (2020), 105759.
[90] T. Wang, T. Gao, H. Guo, Y. Wang, X. Zhou, J. Tian, et al., Preoperative prediction of parametrial invasion in early-stage cervical cancer with MRI-based radiomics nomogram, Eur. Radiol. 30 (6) (2020) 3585–3593.
[91] R. Smith-Bindman, L. Poder, E. Johnson, D.L. Miglioretti, Risk of malignant ovarian cancer based on ultrasonography findings in a large unselected population, JAMA Intern. Med. 179 (1) (2019) 71–77.
[92] R.F. Andreotti, D. Timmerman, L.M. Strachowski, W. Froyman, B.R. Benacerraf, G.L. Bennett, et al., O-RADS US risk stratification and management system: a consensus guideline from the acr ovarian-adnexal reporting and data system committee, Radiology 294 (1) (2020) 168–185.
[93] H. Chen, B.W. Yang, L. Qian, Y.S. Meng, X.H. Bai, X.W. Hong, et al., Deep learning prediction of ovarian malignancy at US compared with O-RADS and expert assessment, Radiology 304 (1) (2022) 106–113.
[94] Y. Zhang, Z. Wang, J. Zhang, C. Wang, Y. Wang, H. Chen, et al., Deep learning model for classifying endometrial lesions, J. Transl. Med 19 (1) (2021) 10.
[95] W.E. Pierson, P.N. Peters, M.T. Chang, L.M. Chen, D.A. Quigley, A. Ashworth, et al., An integrated molecular profile of endometrioid ovarian cancer, Gynecol. Oncol. 157 (1) (2020) 55–61.
[96] K.H. Yu, V. Hu, F. Wang, U.A. Matulonis, G.L. Mutter, J.A. Golden, et al., Deciphering serous ovarian carcinoma histopathology and platinum response by convolutional neural networks, BMC Med 18 (1) (2020) 236.
[97] E. Kawakami, J. Tabata, N. Yanaihara, T. Ishikawa, K. Koseki, Y. Iida, et al., Application of artificial intelligence for preoperative diagnostic and prognostic prediction in epithelial ovarian cancer based on blood biomarkers, Clin. Cancer Res 25 (10) (2019) 3006–3015.
[98] E. Rassy, N. Pavlidis, Progress in refining the clinical management of cancer of unknown primary in the molecular era, Nat. Rev. Clin. Oncol. 17 (9) (2020) 541–554.
[99] M.Y. Lu, T.Y. Chen, D.F.K. Williamson, M. Zhao, M. Shady, J. Lipkova, et al., AI-based pathology predicts origins for cancers of unknown primary, Nature 594 (7861) (2021) 106–110.
[100] A. Penson, N. Camacho, Y. Zheng, A.M. Varghese, H. Al-Ahmadie, P. Razavi, et al., Development of genome-derived tumor type prediction to inform clinical cancer care, JAMA Oncol. 6 (1) (2020) 84–91.
[101] J.K. Grewal, B. Tessier-Cloutier, M. Jones, S. Gakkhar, Y. Ma, R. Moore, et al., Application of a neural network whole transcriptome-based pan-cancer method for diagnosis of primary and metastatic cancers, JAMA Netw. Open 2 (4) (2019), e192597.
[102] Y. Zhao, Z. Pan, S. Namburi, A. Pattison, A. Posner, S. Balachander, et al., CUP-AI-Dx: a tool for inferring cancer tissue of origin and molecular subtype using RNA gene-expression data and artificial intelligence, EBioMedicine 61 (2020), 103030.
[103] L. Nguyen, A. Van Hoeck, E. Cuppen, Machine learning-based tissue of origin classification for cancer of unknown primary diagnostics using genome-wide mutation features, Nat. Commun. 13 (1) (2022) 4013.
[104] M. Haniffa, D. Taylor, S. Linnarsson, B.J. Aronow, G.D. Bader, R.A. Barker, et al., A roadmap for the human developmental cell atlas, Nature 597 (7875) (2021) 196–205.
[105] E. Moiso, A. Farahani, H.D. Marble, A. Hendricks, S. Mildrum, S. Levine, et al., Developmental deconvolution for classification of cancer origin, Cancer Discov. (2022). CD-21-1443.

[106] A. González-Martín, B. Pothuri, I. Vergote, R. DePont Christensen, W. Graybill, M. R. Mirza, et al., Niraparib in patients with newly diagnosed advanced ovarian cancer, N. Engl. J. Med. 381 (25) (2019) 2391–2402.
[107] M.R. Mirza, R.L. Coleman, A. González-Martín, K.N. Moore, N. Colombo, I. Ray-Coquard, et al., The forefront of ovarian cancer therapy: update on PARP inhibitors, Ann. Oncol. 31 (9) (2020) 1148–1159.
[108] I. Ray-Coquard, P. Pautier, S. Pignata, D. Pérol, A. González-Martín, R. Berger, et al., Olaparib plus Bevacizumab as first-line maintenance in ovarian cancer, N. Engl. J. Med. 381 (25) (2019) 2416–2428.
[109] K.S. Tewari, M.W. Sill, R.T. Penson, H. Huang, L.M. Ramondetta, L.M. Landrum, et al., Bevacizumab for advanced cervical cancer: final overall survival and adverse event analysis of a randomised, controlled, open-label, phase 3 trial (Gynecologic Oncology Group 240), Lancet 390 (10103) (2017) 1654–1663.
[110] A.N. Fader, D.M. Roque, E. Siegel, N. Buza, P. Hui, O. Abdelghany, et al., Randomized phase II trial of carboplatin-paclitaxel versus carboplatin-paclitaxel-trastuzumab in uterine serous carcinomas that overexpress human epidermal growth factor receptor 2/neu, J. Clin. Oncol. 36 (20) (2018) 2044–2051.
[111] H. Mahdi, M. Hasipek, Y. Guan, D. Grabowski, H.Q. Al-Sudani, Y. Parker, et al., Dual anti-HER2 therapy in HER2+ uterine and ovarian carcinomas: durable effect with combined therapy, Gynecol. Oncol. 154 (suppl_1) (2019) 85–86.
[112] D.T. Le, J.N. Durham, K.N. Smith, H. Wang, B.R. Bartlett, L.K. Aulakh, et al., Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade, Science 357 (6349) (2017) 409–413.
[113] L. Marcus, S.J. Lemery, P. Keegan, R. Pazdur, FDA Approval Summary: pembrolizumab for the treatment of microsatellite instability-high solid tumors, Clin. Cancer Res 25 (13) (2019) 3753–3758.
[114] Y. Lee, J.H. Park, S. Oh, K. Shin, J. Sun, M. Jung, et al., Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning, Nat. Biomed. Eng. (2022).
[115] A. Kleppe, F. Albregtsen, L. Vlatkovic, M. Pradhan, B. Nielsen, T.S. Hveem, et al., Chromatin organisation and cancer prognosis: a pan-cancer study, Lancet Oncol. 19 (3) (2018) 356–369.
[116] C.-J. Tseng, C.-J. Lu, C.-C. Chang, G.-D. Chen, C. Cheewakriangkrai, Integration of data mining classification techniques and ensemble learning to identify risk factors and diagnose ovarian cancer recurrence, Artif. Intell. Med. 78 (2017) 47–54.
[117] A. Oaknin, T.J. Bosse, C.L. Creutzberg, G. Giornelli, P. Harter, F. Joly, et al., Endometrial cancer: ESMO clinical practice guideline for diagnosis, treatment and follow-up, Ann. Oncol. 33 (9) (2022) 860–877.
[118] A.B. Olawaiye, T.P. Baker, M.K. Washington, D.G. Mutch, The new (Version 9) American Joint Committee on cancer tumor, node, metastasis staging for cervical cancer, CA Cancer J. Clin. 71 (4) (2021) 287–298.
[119] K. Matsuo, H. Machida, R.S. Mandelbaum, I. Konishi, M. Mikami, Validation of the 2018 FIGO cervical cancer staging system, Gynecol. Oncol. 152 (1) (2019) 87–93.
[120] Q. Wu, S. Wang, S. Zhang, M. Wang, Y. Ding, J. Fang, et al., Development of a deep learning model to identify lymph node metastasis on magnetic resonance imaging in patients with cervical cancer, JAMA Netw. Open 3 (7) (2020), e2011625.
[121] H. Kitchener, A.M. Swart, Q. Qian, C. Amos, M.K. Parmar, Efficacy of systematic pelvic lymphadenectomy in endometrial cancer (MRC ASTEC trial): a randomised study, Lancet 373 (9658) (2009) 125–136.
[122] P. Benedetti Panici, S. Basile, F. Maneschi, A. Alberto Lissoni, M. Signorelli, G. Scambia, et al., Systematic pelvic lymphadenectomy vs. no lymphadenectomy in early-stage endometrial carcinoma: randomized clinical trial, J. Natl. Cancer Inst. 100 (23) (2008) 1707–1716.
[123] C. Reijnen, E. Gogou, N.C.M. Visser, H. Engerud, J. Ramjith, L.J.M. van der Putten, et al., Preoperative risk stratification in endometrial cancer (ENDORISK) by a Bayesian network model: a development and validation study, PLoS Med 17 (5) (2020), e1003111.
[124] P. Harter, J. Sehouli, I. Vergote, G. Ferron, A. Reuss, W. Meier, et al., Randomized trial of cytoreductive surgery for relapsed ovarian cancer, N. Engl. J. Med. 385 (23) (2021) 2123–2131.
[125] M.H. Baek, E.Y. Park, H.I. Ha, S.Y. Park, M.C. Lim, C. Fotopoulou, et al., Secondary cytoreductive surgery in platinum-sensitive recurrent ovarian cancer: a meta-analysis, J. Clin. Oncol. 40 (15) (2022) 1659–1670.
[126] R.A. Cowan, A.G.Z. Eriksson, S.M. Jaber, Q. Zhou, A. Iasonos, O. Zivanovic, et al., A comparative analysis of prediction models for complete gross resection in secondary cytoreductive surgery for ovarian cancer, Gynecol. Oncol. 145 (2) (2017) 230–235.
[127] X.M. Fan, J. Zhang, S.H. Niu, K.X. Li, C.Z. Song, Secondary cytoreductive surgery in recurrent epithelial ovarian cancer: a prognostic analysis with 103 cases, Int. J. Surg. 38 (2017) 61–66.
[128] P.G. Rose, J.J. Java, M.A. Morgan, A. Alvarez-Secord, J.P. Kesterson, F. B. Stehman, et al., Disease extent at secondary cytoreductive surgery is predictive of progression-free and overall survival in advanced stage ovarian cancer: an nrg oncology/gynecologic oncology group study, Gynecol. Oncol. 143 (3) (2016) 511–515.
[129] G. Bogani, D. Rossetti, A. Ditto, F. Martinelli, V. Chiappa, L. Mosca, et al., Artificial intelligence weights the importance of factors predicting complete cytoreduction at secondary cytoreductive surgery for recurrent ovarian cancer, J. Gynecol. Oncol. 29 (5) (2018), e66.
[130] A. Laios, A. Gryparis, D. DeJong, R. Hutson, G. Theophilou, C. Leach, Predicting complete cytoreduction for advanced ovarian cancer patients using nearest-neighbor models, J. Ovarian Res. 13 (1) (2020) 117.
[131] S. Lheureux, M. Braunstein, A.M. Oza, Epithelial ovarian cancer: evolution of management in the era of precision medicine, CA Cancer J. Clin. 69 (4) (2019) 280–304.
[132] A. Oseledchyk, M.M. Leitao Jr., J. Konner, R.E. O'Cearbhaill, D. Zamarin, Y. Sonoda, et al., Adjuvant chemotherapy in patients with stage I endometrioid or clear cell ovarian cancer in the platinum era: a Surveillance, Epidemiology, and End Results Cohort Study, 2000-2013, Ann. Oncol. 28 (12) (2017) 2985–2993.
[133] H. Huang, Y.L. Feng, T. Wan, Y.N. Zhang, X.P. Cao, Y.W. Huang, et al., Effectiveness of sequential chemoradiation vs concurrent chemoradiation or radiation alone in adjuvant treatment after hysterectomy for cervical cancer: the STARS Phase 3 randomized clinical trial, JAMA Oncol. 7 (3) (2021) 361–369.
[134] R.A. Brooks, G.F. Fleming, R.R. Lastra, N.K. Lee, J.W. Moroney, C.H. Son, et al., Current recommendations and recent progress in endometrial cancer, CA Cancer J. Clin. 69 (4) (2019) 258–279.
[135] D.P. Mysona, L.K.H. Tran, P.M.H. Tran, P.A. Gehrig, L. Van, Le, S. Ghamande, et al., Clinical calculator predictive of chemotherapy benefit in stage 1A uterine papillary serous cancers, Gynecol. Oncol. 156 (1) (2020) 77–84.
[136] D.C. Gulhan, J.J. Lee, G.E.M. Melloni, I. Cortés-Ciriano, P.J. Park, Detecting the mutational signature of homologous recombination deficiency in clinical samples, Nat. Genet. 51 (5) (2019) 912–919.
[137] H.G. Prigerson, Y. Bao, M.A. Shah, M.E. Paulk, T.W. LeBlanc, B.J. Schneider, et al., Chemotherapy use, performance status, and quality of life at the end of life, JAMA Oncol. 1 (6) (2015) 778–784.
[138] A.A. Elfiky, M.J. Pany, R.B. Parikh, Z. Obermeyer, Development and application of a machine learning approach to assess short-term mortality risk among patients with cancer starting chemotherapy, JAMA Netw. Open 1 (3) (2018), e180926.
[139] C. Sun, X. Tian, Z. Liu, W. Li, P. Li, J. Chen, et al., Radiomic analysis for pretreatment prediction of response to neoadjuvant chemotherapy in locally advanced cervical cancer: a multicentre study, EBioMedicine 46 (2019) 160–169.
[140] U.A. Matulonis, A.K. Sood, L. Fallowfield, B.E. Howitt, J. Sehouli, B.Y. Karlan, Ovarian cancer, Nat. Rev. Dis. Prim. 2 (1) (2016) 16061.
[141] K. Chen, H. Xu, Y. Lei, P. Lio, Y. Li, H. Guo, et al., Integration and interplay of machine learning and bioinformatics approach to identify genetic interaction related to ovarian cancer chemoresistance, Brief. Bioinform 22 (6) (2021).
[142] K.M. Boehm, E.A. Aherne, L. Ellenson, I. Nikolovski, M. Alghamdi, I. Vázquez-García, et al., Multimodal data integration using machine learning improves risk stratification of high-grade serous ovarian cancer, Nat. Cancer 3 (6) (2022) 723–733.
[143] C. Marth, D. Reimer, A.G. Zeimet, Front-line therapy of advanced epithelial ovarian cancer: standard treatment, Ann. Oncol. 28 (suppl_8) (2017) viii36-9.
[144] C.W. Wang, C.C. Chang, Y.C. Lee, Y.J. Lin, S.C. Lo, P.C. Hsu, et al., Weakly supervised deep learning for prediction of treatment effectiveness on ovarian cancer from histopathology images, Comput. Med. Imaging Graph. 99 (2022), 102093.
[145] A.C. Berger, A. Korkut, R.S. Kanchi, A.M. Hegde, W. Lenoir, W. Liu, et al., A comprehensive pan-cancer molecular study of gynecologic and breast cancers, Cancer Cell 33 (4) (2018), 690-705.e9.
[146] Z. Li, S. Chen, W. Feng, Y. Luo, H. Lai, Q. Li, et al., A pan-cancer analysis of HER2 index revealed transcriptional pattern for precise selection of HER2-targeted therapy, EBioMedicine 62 (2020), 103074.
[147] Y.W. Hsiao, C.L. Tao, E.Y. Chuang, T.P. Lu, A risk prediction model of gene signatures in ovarian cancer through bagging of GA-XGBoost models, J. Adv. Res 30 (2021) 113–122.
[148] K. Odunsi, Immunotherapy in ovarian cancer, Ann. Oncol. 28 (suppl_8) (2017) viii1-7.
[149] M. Desbois, A.R. Udyavar, L. Ryner, C. Kozlowski, Y. Guan, M. Dürrbaum, et al., Integrated digital pathology and transcriptome analysis identifies molecular mediators of T-cell exclusion in ovarian cancer, Nat. Commun. 11 (1) (2020) 5583.
[150] L. He, D. Bulanova, J. Oikkonen, A. Häkkinen, K. Zhang, S. Zheng, et al., Network-guided identification of cancer-selective combinatorial therapies in ovarian cancer, Brief. Bioinform 22 (6) (2021).
[151] W.C. Shen, S.W. Chen, K.C. Wu, T.C. Hsieh, J.A. Liang, Y.C. Hung, et al., Prediction of local relapse and distant metastasis in patients with definitive chemoradiotherapy-treated cervical cancer by deep learning from [(^{18})F]-fluorodeoxyglucose positron emission tomography/computed tomography, Eur. Radiol. 29 (12) (2019) 6741–6749.
[152] M. Ferreira, P. Lovinfosse, J. Hermesse, M. Decuypere, C. Rousseau, F. Lucia, et al., [(^{18})F]FDG PET radiomics to predict disease-free survival in cervical cancer: a multi-scanner/center study with external validation, Eur. J. Nucl. Med. Mol. Imaging 48 (11) (2021) 3432–3443.
[153] J.C. Hong, N.C.W. Eclov, N.H. Dalal, S.M. Thomas, S.J. Stephens, M. Malicki, et al., System for high-intensity evaluation during radiation therapy (SHIELD-RT): a prospective randomized study of machine learning-directed clinical evaluations during radiation and chemoradiation, J. Clin. Oncol. 38 (31) (2020) 3652–3661.
[154] K.J. Ramchandran, J.W. Shega, J. Von Roenn, M. Schumacher, E. Szmuilowicz, A. Rademaker, et al., A predictive model to identify hospitalized cancer patients at risk for 30-day mortality based on admission criteria via the electronic medical record, Cancer 119 (11) (2013) 2074–2080.
[155] E.L. Barber, R. Garg, C. Persenaire, M. Simon, Natural language processing with machine learning to predict outcomes after ovarian cancer surgery, Gynecol. Oncol. 160 (1) (2021) 182–186.
[156] R.B. Parikh, C. Manz, C. Chivers, S.H. Regli, J. Braun, M.E. Draugelis, et al., Machine learning approaches to predict 6-month mortality among patients with cancer, JAMA Netw. Open 2 (10) (2019), e1915997.

[157] C.R. Manz, J. Chen, M. Liu, C. Chivers, S.H. Regli, J. Braun, et al., Validation of a machine learning algorithm to predict 180-day mortality for outpatients with cancer, JAMA Oncol. 6 (11) (2020) 1723–1730.
[158] F.J. Zachariah, L.A. Rossi, L.M. Roberts, L.D. Bosserman, Prospective comparison of medical oncologists and a machine learning model to predict 3-month mortality in patients with metastatic solid tumors, JAMA Netw. Open 5 (5) (2022), e2214514.
[159] O. Morin, M. Vallières, S. Braunstein, J.B. Ginart, T. Upadhaya, H.C. Woodruff, et al., An artificial intelligence framework integrating longitudinal electronic health records with real-world data enables continuous pan-cancer prognostication, Nat. Cancer 2 (7) (2021) 709–722.
[160] E. Huynh, A. Hosny, C. Guthier, D.S. Bitterman, S.F. Petit, D.A. Haas-Kogan, et al., Artificial intelligence in radiation oncology, Nat. Rev. Clin. Oncol. 17 (12) (2020) 771–781.
[161] D. Tellez, G. Litjens, P. Bándi, W. Bulten, J.M. Bokhorst, F. Ciompi, et al., Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology, Med. Image Anal. 58 (2019), 101544.
[162] D. Singh, B. Singh, Investigating the impact of data normalization on classification performance, Appl. Soft. Comput. 97 (2020), 105524.
[163] C.-L. Yang, C.-Y. Yang, Z.-X. Chen, N.-W. Lo, Multivariate time series data transformation for convolutional neural network, 2019 IEEE/SICE International Symposium on System Integration (SII), IEEE, 2019, pp. 188–192.
[164] C.J. Kelly, A. Karthikesalingam, M. Suleyman, G. Corrado, D. King, Key challenges for delivering clinical impact with artificial intelligence, BMC Med 17 (1) (2019) 1–9.
[165] J.R. Zech, M.A. Badgeley, M. Liu, A.B. Costa, J.J. Titano, E.K. Oermann, Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study, PLoS Med 15 (11) (2018), e1002683.
[166] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, et al., A guide to deep learning in healthcare, Nat. Med. 25 (1) (2019) 24–29.
[167] D. Kotecha, F.W. Asselbergs, S. Achenbach, S.D. Anker, D. Atar, C. Baigent, et al., CODE-EHR best-practice framework for the use of structured electronic health-care records in clinical research, Lancet Digit. Health 4 (10) (2022) e757–e764.
[168] B. Kang, J. Yoon, H.Y. Kim, S.J. Jo, Y. Lee, H.J. Kam, Deep-learning-based automated terminology mapping in OMOP-CDM, J. Am. Med. Inform. Assoc. 28 (7) (2021) 1489–1496.
[169] R. Makadia, P.B. Ryan, Transforming the premier perspective hospital database into the Observational Medical Outcomes Partnership (OMOP) common data model, EGEMS (Wash., DC) 2 (1) (2014), 1110.
[170] R.B. Ness, Influence of the HIPAA privacy rule on health research, JAMA 298 (18) (2007) 2164–2170.
[171] L. Na, C. Yang, C.-C. Lo, F. Zhao, Y. Fukuoka, A. Aswani, Feasibility of Reidentifying individuals in large national physical activity data sets from which protected health information has been removed with use of machine learning, JAMA Netw. Open 1 (8) (2018), e186040.
[172] K.D. Mandl, E.D. Perakslis, HIPAA and the leak of "deidentified" EHR data, N. Engl. J. Med. 384 (23) (2021) 2171–2173.
[173] W.N. Price 2nd, I.G. Cohen, Privacy in the age of medical big data, Nat. Med. 25 (1) (2019) 37–43.
[174] C.L. Andaur Navarro, J.A.A. Damen, T. Takada, S.W.J. Nijman, P. Dhiman, J. Ma, et al., Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review, BMJ 375 (2021) n2281.
[175] M. Nagendran, Y. Chen, C.A. Lovejoy, A.C. Gordon, M. Komorowski, H. Harvey, et al., Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies, BMJ 368 (2020) m689.
[176] A.J. London, Artificial intelligence and black-box medical decisions: accuracy versus explainability, Hastings Cent. Rep. 49 (1) (2019) 15–21.
[177] G. Novakovsky, N. Dexter, M.W. Libbrecht, W.W. Wasserman, S. Mostafavi, Obtaining genetics insights from deep learning via explainable artificial intelligence, Nat. Rev. Genet. 24 (2) (2023) 125–137.
[178] M. Reyes, R. Meier, S. Pereira, C.A. Silva, F.-M. Dahlweid, Hv Tengg-Kobligk, et al., On the interpretability of artificial intelligence in radiology: challenges and opportunities, Radiol.: Artif. Intell. 2 (3) (2020), e190043.
[179] G. Yang, Q. Ye, J. Xia, Unbox the black-box for the medical explainable AI via multi-modal and multi-centre data fusion: a mini-review, two showcases and beyond, Inf. Fusion. 77 (2022) 29–52.
[180] D. Vale, A. El-Sharif, M. Ali, Explainable artificial intelligence (XAI) post-hoc explainability methods: risks and limitations in non-discrimination law, Ai. Ethics 2 (4) (2022) 815–826.
[181] A. Chattopadhay, A. Sarkar, P. Howlader, V.N. Balasubramanian, Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks, 2018 IEEE winter conference on applications of computer vision (WACV), IEEE, 2018, pp. 839–847.
[182] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?" Explaining the predictions of any classifier, Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining, 2016, pp. 1135–1144.
[183] A. Hosny, C. Parmar, J. Quackenbush, L.H. Schwartz, H.J. Aerts, Artificial intelligence in radiology, Nat. Rev. Cancer 18 (8) (2018) 500–510.
[184] K. Bera, N. Braman, A. Gupta, V. Velcheti, A. Madabhushi, Predicting cancer outcomes with radiomics and artificial intelligence in radiology, Nat. Rev. Clin. Oncol. 19 (2) (2022) 132–146.
[185] E. Begoli, T. Bhattacharya, D. Kusnezov, The need for uncertainty quantification in machine-assisted medical decision making, Nat. Mach. Intell. 1 (1) (2019) 20–23.
[186] Y. Dong, J. Wan, X. Wang, J.H. Xue, J. Zou, H. He, et al., A polarization-imaging-based machine learning framework for quantitative pathological diagnosis of cervical precancerous lesions, IEEE Trans. Med. Imaging 40 (12) (2021) 3728–3738.
[187] A. Iantsen, M. Ferreira, F. Lucia, V. Jaouen, C. Reinhold, P. Bonaffini, et al., Convolutional neural networks for PET functional volume fully automatic segmentation: development and validation in a multi-center setting, Eur. J. Nucl. Med. Mol. Imaging 48 (11) (2021) 3444–3456.
[188] B. Rigaud, B.M. Anderson, Z.H. Yu, M. Gobeli, G. Cazoulat, J. Söderberg, et al., Automatic segmentation using deep learning to enable online dose optimization during adaptive radiation therapy of cervical cancer, Int. J. Radiat. Oncol. Biol. Phys. 109 (4) (2021), 1096-10.
[189] K. Matsuo, S. Purushotham, B. Jiang, R.S. Mandelbaum, T. Takiuchi, Y. Liu, et al., Survival outcome prediction in cervical cancer: Cox models vs deep-learning model, Am. J. Obstet. Gynecol. 220 (4) (2019), 381.e1-14.
[190] H. Sun, X. Zeng, T. Xu, G. Peng, Y. Ma, Computer-aided diagnosis in histopathological images of the endometrium using a convolutional neural network and attention mechanisms, IEEE J. Biomed. Health Inf. 24 (6) (2020) 1664–1676.
[191] F. Christiansen, E.L. Epstein, E. Smedberg, M. Åkerlund, K. Smith, E. Epstein, Ultrasound image analysis using deep neural networks for discriminating between benign and malignant ovarian tumors: comparison with expert subjective assessment, Ultrasound Obstet. Gynecol. 57 (1) (2021) 155–163.
[192] A. BenTaieb, H. Li-Chang, D. Huntsman, G. Hamarneh, A structured latent model for ovarian carcinoma subtyping from histopathology slides, Med. Image Anal. 39 (2017) 194–205.
[193] A.B. Shinagare, P. Balthazar, I.K. Ip, R. Lacson, J. Liu, N. Ramaiya, et al., High-grade serous ovarian cancer: use of machine learning to predict abdominopelvic recurrence on ct on the basis of serial Cancer Antigen 125 levels, J. Am. Coll. Radio. 15 (8) (2018) 1133–1138.