Gynecologic Oncology 177 (2023) 109–116

Contents lists available at ScienceDirect

Gynecologic Oncology

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

Simple rules, O-RADS, ADNEX and SRR model: Single oncologic center validation of diagnostic predictive models alone and combined (two-step strategy) to estimate the risk of malignancy in adnexal masses and ovarian tumors

Giulia Spagnol [a], Matteo Marchetti [a], Orazio De Tommasi [a], Amerigo Vitagliano [b], Francesco Cavallin [c], Roberto Tozzi [a], Carlo Saccardi [a], Marco Noventa [a,*]

[a] Department of Women and Children's Health, Unit of Gynecology and Obstetrics, University of Padua, Padua, Italy
[b] Department of Biomedical and Human Oncological Science (DIMO), 1st Unit of Obstetrics and Gynecology, University of Bari, Policlinico, Bari, Italy
[c] Independent Statistician (collaboration with University of Padua), Solagna, Italy

HIGHLIGHTS

- O-RADS, ADNEX and the two-step strategy (SR plus ADENX) have a sensitivity exceeding 90% for discriminating adnexal masses.
- SR plus ADNEX model applicated to inconclusive SR patients have the highest specificity and positive predictive value.
- The association between SR and ADNEX model applicated to inconclusive SR patients presented the best performance accuracy

ARTICLE INFO

Article history:
Received 16 June 2023
Received in revised form 2 August 2023
Accepted 21 August 2023
Available online 1 September 2023

Keywords:
Ovarian cancer
Simple rules
SRR
ADNEX
O-RADS
Ultrasound

ABSTRACT

Objective. To compare performance of Assessment of Different NEoplasias in the adneXa (ADNEX model), Ovarian-Adnexal Reporting and Data System (O-RADS), Simple Rules Risk (SRR) assessment and the two-step strategy based on the application of Simple Rules (SR) followed by SRR and SR followed by ADNEX in the pre-operative discrimination between benign and malignant adnexal masses (AMs).

Methods. We conducted a retrospective study from January-2018 to December-2021 in which consecutive patients with at AMs were recruited. Accuracy metrics included sensitivity (SE) and specificity (SP) with their 95% confidence intervals (CI) were calculated for ADNEX, O-RADS and SRR. When SR was inconclusive a "two-step strategy" was adopted applying SR + ADNEX model and SR + SRR assessment.

Results. A total of 514 women were included, 400 (77.8%) had a benign ovarian tumor and 114 (22.2%) had a malignant tumor. At a threshold malignancy risk of >10%, the SE and SP of ADNEX model, O-RADS and SRR were: 0.92 (95% CI, 0.86–0.96) and 0.88 (95% CI, 0.85–0.91); 0.93 (95% CI, 0.87–0.97) and 0.89 (95% CI, 0.96–0.92); 0.88 (95% CI, 0.80–0.93) and 0.84 (95% CI, 0.80–0.87), respectively. When we applied SR, 109 (21.2%) cases resulted inconclusive. The SE and SP of two-step strategy SR + SRR assessment and SR + ADNEX model were 0.88 (95% CI, 0.80–0.93) and 0.92 (95% CI, 0.89–0.94), SR + ADNEX model 0.90 (95% CI, 0.83–0.95) and 0.93 (95% CI, 0.90–0.96), respectively.

Conclusions. O-RADS presented the highest SE, similar to ADNEX model and SR + ADNEX model. However, the SR + ADNEX model presented the higher performance accuracy with the higher SP and PPV. This two-step strategy, SR and ADNEX model applicated to inconclusive SR, is convenient for clinical evaluation.

1. Introduction

Adnexal masses (AMs) represent a common finding during a gynecological examination and their differentiation between benign and malignant lesions has high clinical relevance, leading patients with

* Corresponding author at: Department of Women and Children's Health, Unit of Gynecology and Obstetrics, Via Giustiniani 3, 35128 Padova, Italy.
E-mail address: marco.noventa.2@unipd.it (M. Noventa).

https://doi.org/10.1016/j.ygyno.2023.08.012

benign lesion to a conservative approach, while referring patients with suspected ovarian cancer (OC) to tertiary level hospitals for appropriate surgical and medical management [1–3].

Transvaginal-ultrasound (TV-US) by an expert examiner represents the primary imaging modality for preparative assessment of AMs [4,5]. To help clinicians decide on appropriate management, predictive models based on TVS-US have been developed. One of the first and, in the past, most widespread model is the Risk of Malignancy Index (RMI) [6–9]. Subsequently, three prediction models have been developed by the International Ovarian Tumor Analysis (IOTA) group with the aim of discriminating patients with suspicion of malignant versus benign lesions.

In 2008, the IOTA group proposed the use of Simple Rules (SR) to standardize the ultrasonography across different centers, for the diagnosis of OC. This classification predicted whether an AM is likely benign or malignant with approximately 25% of cases falling into an "inconclusive" category. When applicable SR showed a sensitivity of about 93% and a specificity of 95% [10]. Few years later, the Simple Rules Risk (SRR) model, a logistic regression model based on TV-US features of SR, was developed to provide an estimated risk of malignancy for any type of AM [11]. With the aim of increasing diagnostic accuracy (DA), avoiding inconclusive results, IOTA group developed in 2014 the Assessment of Different NEoplasias in the adneXa (ADNEX). This model was applicable to all AMs to predict the general risk of malignancy and to classify the lesion as benign, malignant, borderline ovarian tumor (BOT), OC stage I-IV and metastasis. ADNEX and SRR demonstrated high and comparable sensitivity (approximately 95%), with a lower specificity (approximately 70%) [12]. Although the good DA of these models, their acceptance has been limited in clinical practice in some countries such as the United States. More recently, the American College of Radiology (ACR) published the Ovarian-Adnexal Reporting and Data System (O-RADS) to stratify risk of malignancy of AM and to propose the management recommendations in each O-RADS risk category, improving sonographic interpretation [13].

In 2021 the first Consensus Statement on pre-operative diagnosis of ovarian tumors was published by ESGO/ISUOG/IOTA/ESGE, outlining the significance of using these models in the pre-operative assessment of AMs [14].

Considering the reported DA of SRR and ADNEX, their application as first line approach for AMs could result in a high number of false positive [15]. So, in clinical practice, when subjective assessment (SA) by an expert is not available, an approach that can reduce the percentage of false positives by improving the specificity of the test would be desirable. Starting from this assumption, we investigated the application of SR as first-line approach. If not applicable, a mathematical model has been used to estimate the risk of malignancy. To the best of our knowledge, this two-step strategy has not been validated before. The aim of our study was to compare, for the first time, the performance of the ADNEX model, O-RADS, SRR and the two-step strategy based on the application of SR followed by SRR and SR followed by the ADNEX model in the pre-operative discrimination between benign and malignant AMs.

2. Methods

2.1. Study design

This was a retrospective, diagnostic test accuracy (DTA) study on pre-operative discrimination between benign and malignant AMs. We included all consecutive patients with AM that undergone surgery at the Unit of Gynecology and Obstetrics of University of Padova from Jan-2018 to Dec-2021.

The study assessed four models and their association for the preoperative evaluation of AMs: SR; ADNEX, SRR and O-RADS. When SR resulted inconclusive, we applied to these patients the SRR and ADNEX model to calculate and compare their associated performance. Histopathological diagnosis after surgical removal of the lesion was the reference standard (evaluated results: benign; BOT; OC). All excised tissues were examined histologically in our surgical pathology department following the guidelines of the World Health Organization (WHO) [16].

All patients signed a document approved by our institution for anonymous use according to European privacy law. The institutional review board approved the study (IRB: 44n/AO/2020).

2.2. Inclusion and exclusion criteria

The inclusion criteria were: i) patients who underwent surgical procedure at our centre for an AM ii) histopathological result available; iii) patients who underwent TVS-US and CA125 serum assay a maximum of 3 months before the surgical procedure; iv) description of AM with IOTA definition, result of ADNEX model, O-RADS and SRR for each patient included calculated before surgery; v)18 years of age or older.

Exclusion criteria were as follows: (1) patients' carrier of genetic mutations/syndromes, as BRCA1 and 2 and Lynch II syndrome; (2) previous confirmed diagnosis of OC; (3) pregnancy state; (4) >3 months elapsed among TVS-US and CA125 serum assay and surgical procedure.

2.3. Data collection

Patients were identified through our institution computer database. For each patient the investigators reviewed the electronic hospital records and pathology reports to determine study eligibility, patients' general features, histological report, TV-US information.

TV-US was the primary approach; also, we performed transabdominal ultrasound when the lesion was too large. TV-US was performed by two IOTA-certified gynaecologic specialist or a gynaecologist in training under supervision. All data from ultrasound were obtained from original reports using our institution computer database initiated to collect clinical information at point of care. All IOTA variables, useful to calculate SR, ADNEX and SRR, were collected at time of ultrasound. All three models were calculated at time of ultrasound or before surgery, so all operators were blinded to histopathology result. Patients with missing information were excluded. The ultrasound machines used were Voluson E8 and S10 (GE Healthcare, Zipf, Austria), with 5.0–9.0 MHz TV-probes and 1.0–5.0 MHz transabdominal probes.

All women received specific surgical treatment at our clinic according to the age and the suspect of adnexal pathology. After surgery, the tissue was examined by an experienced gynaecologic pathologist for histopathology and grade. For each patient we collected the following data: i) age, ii) menopausal status: premenopausal (pre-M) and postmenopausal (post-M), iii) histopathology features, iv) CA125, v) the results of ADNEX model, O-RADS, and SRR.

2.4. Predictive models

i) IOTA SR: the IOTA SR included ten ultrasound rules: five ultrasound rules that are classified as benign and five that are considered malignant. The benign and malignant features are reassumed in Timmerman et al. [10]. If both the B and M features were either present or absent the SR classified the lesion as "inconclusive". (http://www.iotagroup.org).

When the SR resulted inconclusive, we applied ADNEX model (See ADNEX model paragraph) and SRR (See SRR paragraph).

ii) ADNEX model: ADNEX model considers nine variables and the formula for the risk calculation can be found in the original article Van Calster et al. [12]. The result of the model is an absolute percentage risk to discriminate between benign and malignant lesion. For the clinical practice an application is available at the website: http://www.iotagroup.org/adnexmodel.

iii) O-RADS: O-RADS, proposed by ACR, relies on the ultrasound nomenclature developed by the IOTA group. The O-RADS US working

group defined six categories for risk classification: O-RADS 0, an incomplete evaluation; O-RADS 1, normal sonographic ovarian morphology; O-RADS 2, the almost certainly benign category (risk of malignancy <1%); O-RADS 3, lesions with low risk of malignancy (1%–10%); O-RADS 4, lesions with intermediate risk of malignancy (10%–50%); O-RADS 5, lesions with high risk of malignancy (>50%).

iv) SRR: the SRR assessment utilizes the same five benign and the five malignant rules used in SR. The presence or absence of each rule is entered into the Simple Rules Risk calculator available on-line (https://homes.esat.kuleuven.be/~sistawww/biomed/ssrisk/). The calculator results in a numeric risk of malignancy based on combinate ultrasound characteristics of the AMs and type of center (oncology vs non-oncology center) thus eliminating the inconclusive classification.

2.5. Statistical analysis

Descriptive variables were summarized as median and interquartile range (IQR) (continuous variables), or frequency and percentage (categorical variables). Accuracy analysis compared all malignancies (OC and BOT lesions) vs. benign lesions in all women, and in pre-M and post-M women. Accuracy metrics included sensitivity (SE), specificity (SP), positive predictive value (PPV) and negative predictive value (NPV) with their 95% confidence intervals (CI). Accuracy metrics were calculated for ADNEX, O-RADS and SRR using the recommended 10% cut-off. When assessing the accuracy of SR + ADNEX and SR + SRR, a "two-step strategy" (applying IOTA SR as first step and using ADNEX/SRR as second step if IOTA SR was indeterminate) was adopted. In addition, receiver operating characteristic (ROC) curve was plotted to illustrate diagnostic ability and optimal thresholds for ADNEX and SRR. Statistical analysis was performed using R 4.3 (R Foundation for Statistical Computing, Vienna, Austria) [17].

3. Results

3.1. Patient and tumor characteristics

Over the study period a total of 514 patients resulted eligible according to our inclusion and exclusion criteria (Fig. 1). Patient characteristics are summarized in Table 1. We collected data about 260 (51.6%) pre-M and 254 (49.4%) post-M women. Histopathology revealed 89 (17.3%) OC, 25 (4.9%) BOTs and 400 (77.8%) benign tumors (Table 2).

SR were informative in 405 (78.8%) cases (345 benign and 60 malignant) and the results were inconclusive in 109 (21.2%) cases. The histological diagnosis of the 109 inconclusive cases were 56 benign lesions (51.4%) and 53 malignant lesions (48.6%) (Table 3). In the 405 cases where the SR could be applied, in all women the SE and SP were 0.89 (0.96 to 0.92) and 0.98 (0.89 to 0.98) respectively, and in post-M were 0.93 and 0.97 respectively.

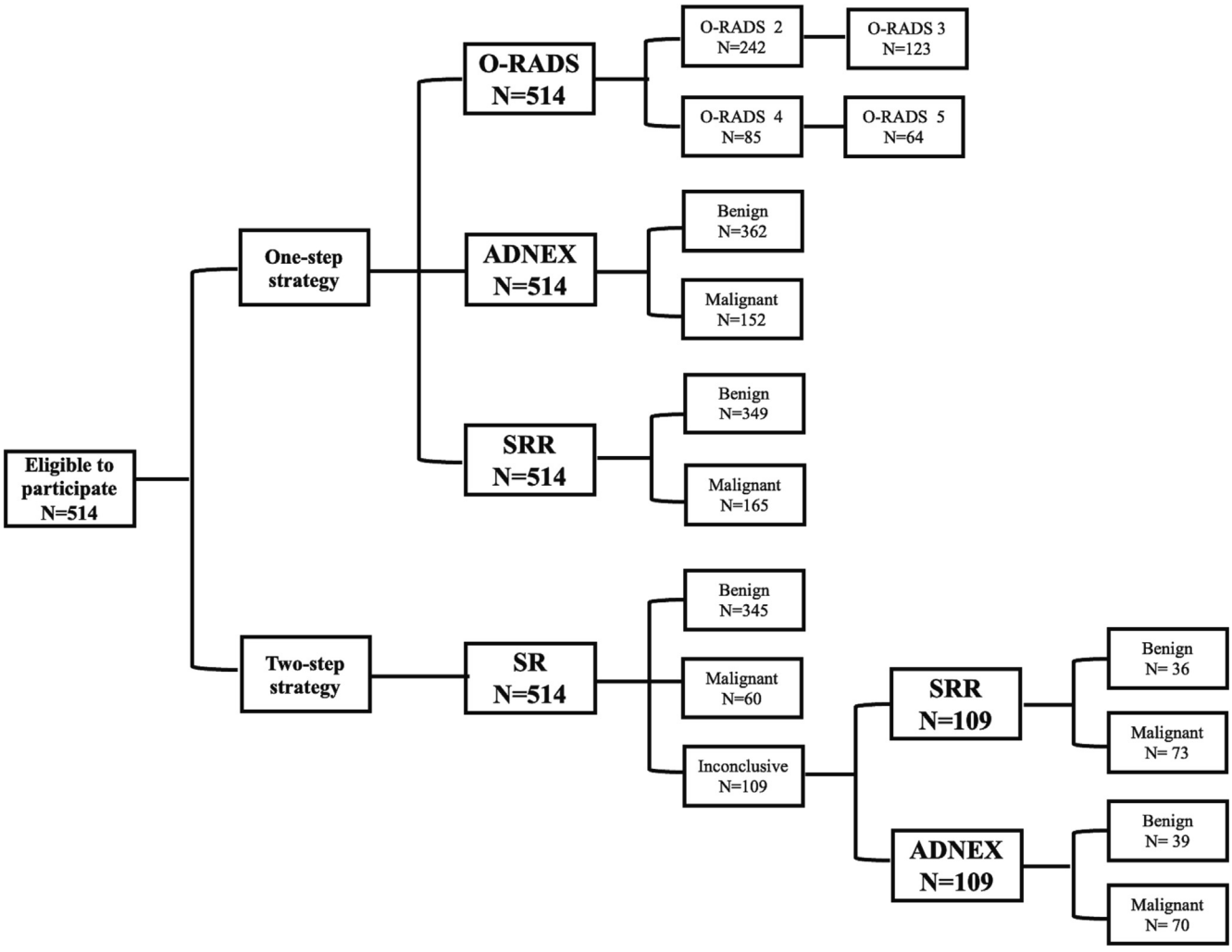

Fig. 1. Flow diagram of our study. AM = adnexal masses; O-RADS = Ovarian-Adnexal Reporting and Data System; Reporting and Data System; IOTA = International Ovarian Tumor Analysis; SR = simple rules; SRR simple rules risk.

Table 1
Patient characteristics.

	All women				Premenopausal women			Postmenopausal women		
	Total	Benign	Malign	BOT	Benign	Malign	BOT	Benign	Malign	BOT
N	514	400	89	25	218	27	15	182	62	10
Age, years [a]	51 (41–63)	49 (38–62)	58 (49–70)	46 (34–57)	40 (33–46)	45 (42–47)	39 (31–45)	63 (56–71)	63 (58–72)	59 (57–69)
Pre-M	260 (51%)	–	–	–	–	–	–	–	–	–
Post-M	254 (49%)									
SR:										
Indeterminate	109 (21%)	56 (14%)	33 (37%)	20 (80%)	27 (12%)	11 (41%)	13 (86%)	29 (16%)	22 (35%)	7 (70%)
Benign	345 (67%)	338 (85%)	4 (5%)	3 (12%)	190 (87%)	3 (11%)	1 (7%)	148 (81%)	1 (2%)	2 (20%)
Malign	60 (12%)	6 (1%)	52 (58%)	2 (8%)	1 (1%)	13 (48%)	1 (7%)	5 (3%)	39 (63%)	1 (10%)
ADNEX [a]	3.9 (2.6–20.1)	3.2 (2.4–4.8)	83.4 (52.6–97.0)	27.0 (15.1–58.9)	2.9 (2.2–4.1)	69.6 (37.1–84.8)	27.0 (22.2–39.3)	3.8 (2.7–5.8)	90.3 (58.5–97.7)	39.0 (5.2–78.2)
O-RADS [a]	3 (2–4)	2 (2–3)	5 (4–5)	4 (4–4)	2 (2–3)	4 (4–5)	4 (4–4)	3 (2–3)	5 (4–5)	4 (3–4)
SRR [a]	3.1 (0.5–15.2)	2.9 (0.5–3.1)	89.5 (71.7–97.6)	23.4 (3.1–71.7)	2.7 (0.5–3.1)	81.7 (64.3–90.8)	48.7 (12.8–69.7)	3.1 (0.5–5.7)	89.5 (71.7–97.6)	4.4 (2.8–63.9)

Legend: SR: simple rules; ADNEX: Assessment of Different NEoplasias in the adnexa; O-RADS: Ovarian-Adnexal Reporting and Data System; SRR: Simple Rules risk; post-M: post-menopause; pre-M: pre-menopause.
Data summarized as n (%) or [a] median (IQR).

3.2. One-step strategy (predictive models alone)

3.2.1. ADNEX model

According to ADNEX model we classified as benign 362 AMs and as malignant 152 AMs (Fig. 1).

The percentages of malignancy at final histology in benign and malignant AMs according to ADNEX model were 2.5% and 69% respectively (Table 3 for details on pre-M and post-M patients).

Setting the model to a cut off of 10% we found for all sample a SE of 0.92 (0.86 to 0.96), SP of 0.88 (0.85 to 0.91), PPV of 0.69 (0.61 to 0.76) and NPV of 0.98 (0.95 to 0.99). Analyzing separately pre-M and post-M women, the performance of ADNEX model was maintained with only a slight decrease and increase of SE in pre-M and post-M patients respectively (Table 4).

The AUC of ADNEX for differentiating between benign and malignant AMs at the time of ultrasound examination was 0.94. This high value was maintained also considering only pre-M and post-M patients (AUC 0.92 and 0.95 respectively). In our sample, ROC curves suggested that the ADNEX model optimal threshold for discriminating benign and malignant AMs was 14.8 for all women, 12.5 for pre-M and 26.7 for post-M women (Fig. 2a).

Table 2
Histological characteristics of the AM.

Total	514
Benign masses	400 (77.8%)
Theca lutein cyst	45 (11.3%)
Follicular cyst >3 cm	75 (18.8%)
Hemorrhagic cyst	61 (15.3%)
Endometrioma	85 (21.3%)
Mucinous cystadenoma	25 (6.3%)
Dermoid cysts	51 (12.7%)
Serous cystadenoma	32 (8%)
Paraovarian cyst	26 (6.5%)
Malignant masses	114 (22.2%)
HGSOC	56 (49.1%)
LGSOC	3 (2.6%)
Clear cell carcinoma	3 (2.6%)
Endometrioid carcinoma	10 (8.8%)
Mucinous cystadenocarcinoma	4 (3.5%)
Malignant Mullerian tube mixed tumor	3 (2.6%)
Immatur teratoma	2 (1.8%)
Mixed carcinoma	1 (0.9%)
Granular cell tumor	1 (0.9%)
Metastatic carcinoma	6 (5.3%)
Borderline mucinous tumor	16 (14%)
Borderline serous tumor	9 (7.9%)

Legend: AM: adnexal mass; HGSOC: high grade serous ovarian carcinoma; LGSOC: low grade serous ovarian carcinoma; CCC: Clear cell carcinoma.

3.2.2. O-RADS

According to O-RADS we classified as category 2–3, 365 AMs and as category 4–5, 149 AMs (Fig. 1).

The percentages of malignancy at final histology in categories 2–3 and 4–5 were 2.2% and 71.1% respectively. (Table 3 for details on pre-M and post-M patients).

Setting the model to a cut off of 10% (O-RADS 4–5) we found for all women a SE of 0.93 (0.87 to 0.97), SP of 0.89 (0.96 to 0.92), PPV of 0.72 (0.64 to 0.79) and NPV of 0.98 (0.96 to 0.99). Analyzing separately pre-M and post-M women, the performance of O-RADS was maintained with only a slight decrease and increase of SE in pre-M and post-M patients respectively (Table 4).

3.2.3. SRR model

According to SRR model we classified as benign 349 AMs and as malignant 165 AMs (See Fig. 1). The percentages of malignancy at final histology in benign and malignant AMs according to SRR model were 4% and 60% respectively (Table 3 for details on pre-M and post-M patients).

Setting the model to a cut off of 10% we found for all women a SE of 0.88 (0.80 to 0.93), SP of 0.84 (0.80 to 0.87), PPV of 0.61 (0.53 to 0.68) and NPV of 0.96 (0.93 to 0.98). Analyzing separately pre-M and post-M women, the performance of SRR model was maintained with only a slight decrease of SE in pre-M compared to post-M patients (Table 4).

The AUC of SRR for differentiating between benign and malignant adnexal masses at the time of ultrasound examination was 0.92. This high value was maintained also considering only pre-M and post-M patients (AUC 0.91 and 0.92 respectively).

In our sample, ROC curves suggested that the SRR optimal threshold for discriminating benign and malignant AMs was 21.6 for all women, 12.5 for pre-menopausal and 21.6 for postmenopausal women (Fig. 2b).

3.3. Two-step strategy (predictive models combined)

3.3.1. SR + SRR model

Applying SRR model to the inconclusive SR group we classified as benign 36 AMs and as malignant 73 AMs (Fig. 1). The percentages of malignancy at final histology in benign and malignant AMs were 19.4% and 63.0% respectively (Table 3 for details on pre-M and post-M patients).

For all sample, the accuracy metrics of SR + SRR for risk of malignancy showed a SE of 0.88 (0.80 to 0.93), SP of 0.92 (0.89 to 0.94), PPV of 0.75 (0.67 to 0.82) and NPV of 0.96 (0.94 to 0.98). In post-M

Table 3
Malignancy rates for SR, ADNEX model, O-RADS in all patients and pre-M and post-M patients.

	Histology		Malignancy rate (%)
	Benign	Malignant	
All patients			
SR			
Benign: 345	338	7	2.0%
Malignant: 60	6	54	90.0%
Inconclusive: 109	56	53	49%
ADNEX			
Benign: 362	353	9	2.5%
Malignant: 152	47	105	69%
O-RADS			
O-RADS 2–3: 365	357	8	2.2%
O-RADS 4–5: 149	43	106	71.1%
SRR			
Benign: 349	335	14	4.0%
Malignant: 165	65	100	60.0%
Only SR inconclusive			
ADNEX			
Benign: 39	35	4	10.2%
Malignant:70	21	49	70.0%
SRR			
Benign = 36	29	7	19.4%
Malignant = 73	27	46	63.0%
Pre-M patients			
SR			
Benign: 194	190	4	2.1%
Malignant: 15	1	14	93.3%
Inconclusive: 51	27	24	47.1%
ADNEX			
Benign = 206	202	4	1.9%
Malignant = 54	16	38	70.4%
O-RADS			
O-RADS 2–3: 207	203	4	1.9%
O-RADS 4–5: 53	15	38	71.7%
SRR			
Benign: 200	194	6	3.0%
Malignant: 60	24	36	60.0%
Only SR inconclusive			
ADNEX			
Benign = 18	17	1	5.6%
Malignant = 33	10	23	69.7%
SRR			
Benign = 18	16	2	11.1%
Malignant = 33	11	22	66.7%
Post-M patients			
SR			
Benign: 151	148	3	1.9%
Malignant: 45	5	40	88.9%
ADNEX			
Benign = 156	151	5	3.2%
Malignant = 98	31	67	68.3%
O-RADS			
O-RADS 2–3: 158	154	4	2.5%
O-RADS 4–5: 96	28	68	70.8%
SRR			
Benign: 150	142	8	5.3%
Malignant: 104	40	64	61.5%
Only SR inconclusive			
ADNEX			
Benign = 21	18	3	14.3%
Malignant = 37	11	26	70.3%
SRR			
Benign = 18	13	5	27.8%
Malignant = 40	16	24	60.0%

Legend: SR: simple rules; ADNEX: Assessment of Different NEoplasias in the adnexa; O-RADS: Ovarian-Adnexal Reporting and Data System; SRR: Simple Rules risk; post-M: post menopause

patients we found a slight increase of SE 0.89 (0.79 to 0.95) and a slight decrease of SP 0.88 (0.83–0.93); other metrics were stable. The two-step strategy based on the application of SR + SRR showed the lowest performance in terms of SE and intermediate results in terms of SP and PPV compared to other techniques. This trend was confirmed separating data of post-M patients (Table 4).

3.3.2. SR + ADNEX model

Applying ADNEX model to the inconclusive SR group we classified as benign 36 AMs and as malignant 70 AMs (Fig. 1). The percentages of malignancy at final histology in benign and malignant AMs were 10.2% and 70.0% respectively (Table 3 for details on pre-M and post-M patients).

For all sample, the accuracy metrics of SR + ADNEX model for risk of malignancy showed a SE of 0.90 (0.83 to 0.95), SP of 0.93 (0.90 to 0.96), PPV of 0.79 (0.71 to 0.86) and NPV of 0.97 (0.95 to 0.99). In post-M patients we found a slight increase of SE 0.92 (0.83 to 0.97) and a slight decrease of SP 0.91 (0.86 to 0.95); other metrics were stable. The two-step strategy based on the application SR + ADNEX model showed the best performance in terms of SP and PPV and an intermediate result in terms of SE compared to other techniques. This trend was confirmed separating data of post-M patients (Table 4).

4. Discussion

4.1. Rationale

In 2021 ESGO/ISUOG/IOTA/ESGE consensus Statement on pre-operative diagnosis of OC underlined the importance of TV-US examination as the standard first-line imaging investigation for the pre-operative assessment of AMs [14]. Unfortunately, the majority of OC are diagnosed in advanced stages and the guidelines recommended that a gynecologic oncologist performed the appropriate pre-operative assessment and surgery [18,19].

In particular, the guidelines underline that the SA by a clinician with an experience in gynecologic oncology US is the best approach to discriminate between benign and malignant AMs [20–22]. Starting from this assumption, IOTA group and ACR proposed different models (SRR, ADNEX, O-RADS) to discriminate between benign and malignant AMs that could be applied also by clinicians with less experience in oncological diagnostics [11–13]. The fundamental characteristic of these models is that they are always applicable; on the contrary the previous US based diagnostic algorithms (SR and Simple descriptors-SD) lead to a certain percentage of inconclusive results [10,23]. Despite SRR and ADNEX demonstrated a very good diagnostic accuracy [12], it resulted slightly lower when compared to SA [21]. In particular ADNEX, SRR and O-RADS demonstrated high SE for diagnosis of malignant AMs with lower SP [24].

To overcome this potentially high false positive rate stepwise strategies have been proposed. Three studies validated the IOTA three step strategy using SD followed by SR and SA showing excellent diagnostic accuracy [22–25]. However, this model does not exclude the subjective criterion with the above-mentioned limits. Recently, Landofo et al. proposed a two-step strategy based on SD and ADNEX model when benign-SD where not applicable [15]. The AUCs for this two-step strategy was good (0.95 with CA-125). Setting ADNEX at 10% the authors found a SE of 0.91 and SP of 0.85 for all women with a decrease of SP to 0.78 considering post-M patients. Despite the good metrics reported, the SE and SP of the two-step strategy is not much different applying ADNEX as the first method [15]. This result could be attributed to the fact that benign-SD were applicable in 37% of the sample, as reported also by other papers [26]. For this reason, we believe that the application of the SR as first line approach instead of SD could improve diagnostic metrics [27]. In our population the rate of indeterminate results by SR were 21% and considering only benign-SR they resulted applicable in 64% of the sample.

4.2. Synthesis of results and comparison with existing literature

In our population, O-RADS showed the best SE to identify malignant lesions, SE was 0.93 in all women and 0.94 in post-M women; these data was similarly to ADNEX model and SR + ADNEX model that performed

Table 4
Accuracy metrics (sensitivity, specificity, positive predictive value, negative predictive value) for indicators comparing benign disease vs. borderline/malign tumors in all women, and in pre- and postmenopausal women.

Patients	Indicator	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
One-step strategy					
All women	ORADS (>10%)	0.93 (0.87 to 0.97)	0.89 (0.96 to 0.92)	0.72 (0.64 to 0.79)	0.98 (0.96 to 0.99)
	ADNEX (>10%)	0.92 (0.86 to 0.96)	0.88 (0.85 to 0.91)	0.69 (0.61 to 0.76)	0.98 (0.95 to 0.99)
	SRR (>10%)	0.88 (0.80 to 0.93)	0.84 (0.80 to 0.87)	0.61 (0.53 to 0.68)	0.96 (0.93 to 0.98)
Pre-M women	ORADS (>10%)	0.90 (0.77 to 0.97)	0.93 (0.89 to 0.96)	0.72 (0.58 to 0.83)	0.98 (0.95 to 0.99)
	ADNEX (>10%)	0.90 (0.77 to 0.97)	0.93 (0.88 to 0.96)	0.70 (0.56 to 0.82)	0.98 (0.95 to 0.99)
	SRR (>10%)	0.86 (0.71 to 0.95)	0.89 (0.84 to 0.93)	0.60 (0.47 to 0.72)	0.97 (0.93 to 0.99)
Post-M women	ORADS (>10%)	0.94 (0.86 to 0.98)	0.85 (0.79 to 0.90)	0.72 (0.61 to 0.80)	0.97 (0.94 to 0.99)
	ADNEX (>10%)	0.93 (0.85 to 0.98)	0.83 (0.77 to 0.88)	0.68 (0.58 to 0.77)	0.97 (0.93 to 0.99)
	SRR (>10%)	0.89 (0.79 to 0.95)	0.78 (0.71 to 0.84)	0.61 (0.51 to 0.71)	0.95 (0.90 to 0.98)
Two-step strategy					
All women	SR + ADNEX[a]	0.90 (0.83 to 0.95)	0.93 (0.90 to 0.96)	0.79 (0.71 to 0.86)	0.97 (0.95 to 0.99)
	SR + SRR[b]	0.88 (0.80 to 0.93)	0.92 (0.89 to 0.94)	0.75 (0.67 to 0.82)	0.96 (0.94 to 0.98)
Pre-M women	SR + ADNEX[a]	0.88 (0.74 to 0.96)	0.95 (0.91 to 0.97)	0.77 (0.63 to 0.88)	0.98 (0.95 to 0.99)
	SR + SRR[b]	0.86 (0.71 to 0.95)	0.94 (0.91 to 0.97)	0.75 (0.60 to 0.86)	0.97 (0.94 to 0.99)
Post-M women	SR + ADNEX[a]	0.92 (0.83 to 0.97)	0.91 (0.86 to 0.95)	0.80 (0.70 to 0.88)	0.97 (0.93 to 0.99)
	SR + SRR[b]	0.89 (0.79 to 0.95)	0.88 (0.83 to 0.93)	0.75 (0.64 to 0.84)	0.95 (0.91 to 0.98)

Legend: SR: simple rules; ADNEX: Assessment of Different NEoplasias in the adnexa; O-RADS: Ovarian-Adnexal Reporting and Data System; SRR: Simple Rules risk; post-M: post-menopause; pre-M: pre-menopause; PPV: positive predictive value; NPV: negative predictive value

[a] The index adopts SR decision rule of benignancy/malignancy, or ADNEX decision rule if SR is indeterminate.

[b] The index adopts SR decision rule of benignancy/malignancy, or SRR decision rule if SR is indeterminate.

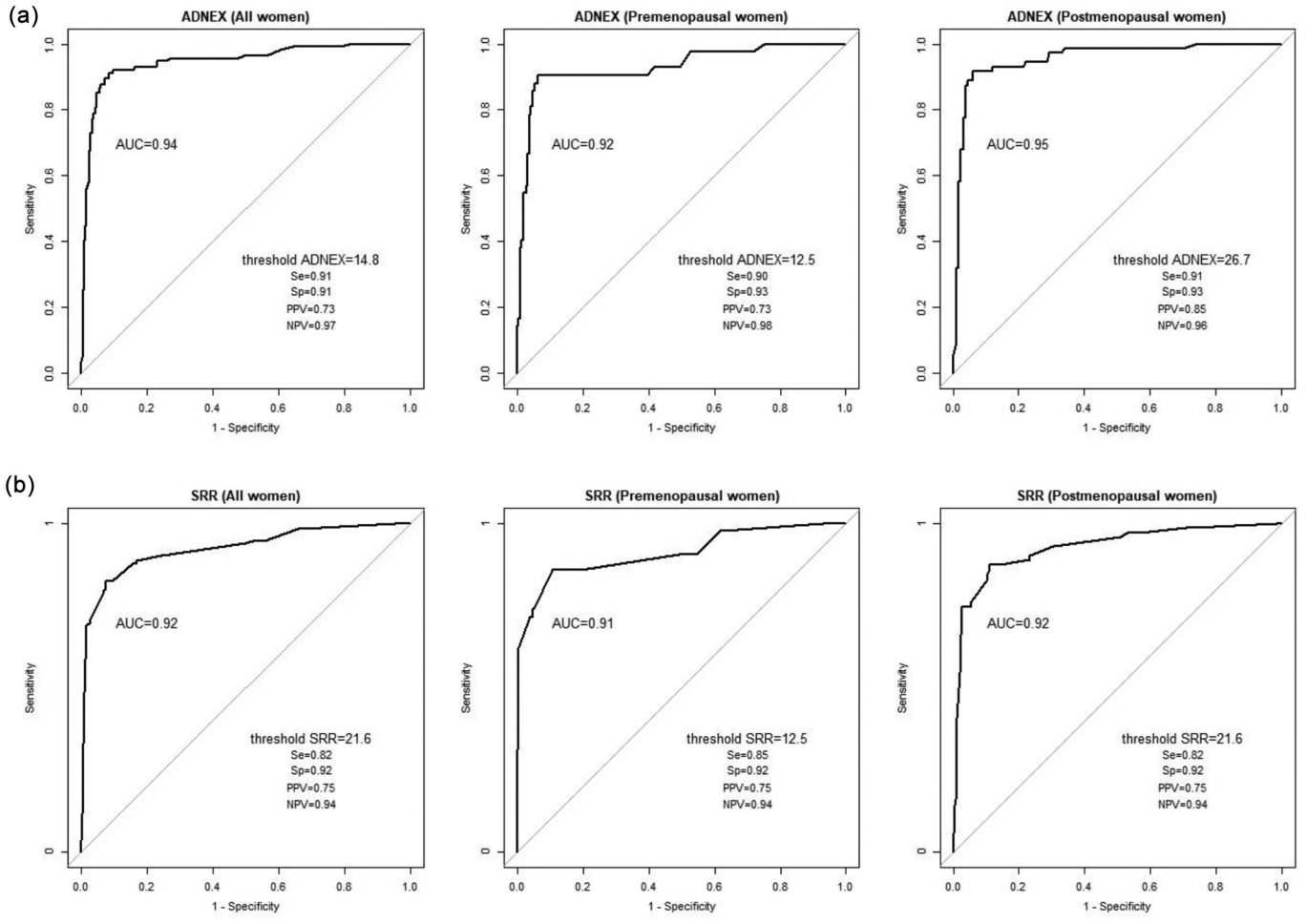

Fig. 2. a. ROC curves for ADNEX predicting malign/borderline tumors in all women, and in pre- and postmenopausal women. b. ROC curves for SRR predicting malign/borderline tumors in all women, and in pre- and postmenopausal women.

well in identifying lesions likely to be malignant with a SE of 0.92 and 0.90 in all women respectively and 0.93 and 0.92 in post-M women; only SRR alone and SR + SRR presented a lower SE (both 0.88 in all women and 0.89 in post-M women).

The AUC of ADNEX model alone was 0.94 when CA 125 was included in the risk calculation that resulted comparable to the recent paper by Landofo et al. [15] The AUC of SRR model alone resulted lower than ADNEX (0.92).

For all women the SP was 0.88 for ADNEX alone, 0.89 for O-RADS and 0.93 applying SR + ADNEX model; for post-M patients the SP was 0.83 for ADNEX alone, 0.85 for O-RADS and 0.91 for SR + ADNEX model. Based on our results, for all women, SR + ADNEX model showed the best accuracy compared to SR + SRR, ADNEX and O-RADS. SR + ADNEX showed a very good SE of 0.90, only slightly lower to O-RADS (0.93) and ADNEX model alone (0.92); the best SP of 0.93 (O-RADS 0.89, ADNEX alone 0.88; SRR alone 0.84 and SR + SRR 0.92). Finally, SR + ADNEX showed the best PPV of 0.79 for all women and 0.80 for post-M women; PPV were found to be particularly higher compared with O-RADS, ADNEX and SRR alone.

Post-test probability PPV is strictly related to SE, SP and prevalence of the disease; an increase of the prevalence leads to an increase of PPV. In our population the prevalence of malignancy was 22.2% that is similar to that reported by other papers, generally between 20 and 30% [31–33]. So, the good metrics of our post-test probability resulted reliable and comparable to literature.

Our results agree with the performance of IOTA models reported in literature, both in larger and smaller studies. The study of validation of the ADNEX model on 5909 cases presented an AUC of 0.94 when CA 125 was not included in the risk calculation [12]. Chen et al. (in 278 women) reported an AUC of 0.94 when CA 125 was included [28]. In these two important studies the performance of the ADNEX model was similar with our results, independent by CA 125 and sample size, with a good performance accuracy. In fact, the inclusion of CA 125 presented a limited impact on the performance of the ADNEX model [29].

Instead, given the recent introduction of O-RADS, data on its performance accuracy and comparative studies were limited. In comparison to the ADNEX model, O-RADS does not require specialized software to determine the risk of malignancy and offers recommendations for additional evaluations such as magnetic resonance imaging (MRI) [13]. In 2021 Cao et al. retrospectively validated the diagnostic performance of O-RADS for the diagnosis of OC. In this study the ROC-analysis using O-RADS 4–5 demonstrated a SE and SP of 98.7% and 83.2%, respectively. The authors concluded that O-RADS US provided an effective malignancy risk stratification for AMs with high reliability for radiologists with different experiences [30].

Differently from this non-comparative study, two important studies compared the US models. In a database of 499 AM, Basha et al. demonstrated a SE and SP of 96.6% and 92.8% for O-RADS and a SE e SP of 92.1% and 93.2% for SR, respectively. The study concluded that O-RADS had higher sensitivity than IOTA SR with relatively similar specificity [31]. After that, in 2022 a study about 150 women was published to compare ADNEX model, SR, SRR assessment and O-RADS. The SE and SP of O-RADS were 100% and 46.4%; ADNEX model were 97.5% and 63.6%; SRR model were 100% and 51.8%, respectively. The authors concluded that IOTA models and O-RADS have similar SE in the discrimination of malignant AM, however IOTA models have higher SP [32].

Recently, Timmerman et al. published a retrospective external validation study using data from IOTA5 to investigate the diagnostic performance of the O-RADS lexicon and the IOTA 2-step strategy: at the 10% risk threshold (O-RADS 4–5), the O-RADS lexicon had 92% SE and 80% SP, and the IOTA 2-step strategy had 91% SE and 85% SP. The findings of this study suggest that both the O-RADS lexicon and the IOTA 2-step strategy can be used to stratify patients into risk groups [33].

Our results were in agreement with these studies. In particular the SE of O-RADS and ADNEX model were similar, instead the SP was highest in our study with a comparable prevalence of cancer (20–30%).

The data about SR + SRR and SR + ADNEX compared with other models were not evaluated in literature so our originally results were not comparable with other studies.

4.3. Strengths and limitations

Main points of strength of our study are related to rigorous data collection methodology and strict inclusion criteria; all information was collected from our electronic hospital records, which are compiled by clinicians at each step of patient's treatment, representing certainly a guarantee of the completeness and correctness of the data reported. We included exclusively patients who had the whole pathway, from the diagnosis to treatment, at our Institution; excluding any sources of bias related to heterogeneous diagnostic and surgical choices. We presented a good sample size with all surgeries performed within 3 months from US examination. All sonographers were experienced and IOTA certified.

Certainly, our study is affected by different limitations. The first is related to the retrospective design and data analysis; however, all data about US evaluation were consecutively reported in our database and electronic hospital record at point of care, thus limiting potential selection bias. Our data were collected in a single oncologic center, which may limit broad application of the findings.

5. Conclusions

This is the first study that proposed and validated the application of the two-step strategy based firstly on SR and subsequently on SRR and ADNEX (with cut-off 10%) for the risk stratification of AMs. Moreover, this one of the first study with large population that compared together the diagnostic accuracy of the main ultrasound-based model (ADNEX, O-RADS, SRR) for the evaluation of AMs. O-RADS presented the highest SE, similar to SE of ADNEX model and SR + ADNEX model, all exceeding 90% in the preoperative discrimination of malignant and benign AM. However, the SR + ADNEX model have higher SP and PPV.

Considering our results, the diagnostic performance of the two-step strategy was better than that of the one-step strategy for the following reason: very similar and comparable SE but with a higher SP and post-test probability PPV. This last metric is particularly important due to the possibility to recognize precisely the malignancies within the test positive, avoiding unnecessary referrals to tertiary gynecological cancer center and unnecessary surgeries. So, as practical recommendations, we suggest to use TV-US two-step strategy firstly applying SR. In cases of malignant results, we suggest immediate referral to an oncological unit; in cases of inconclusive results, we suggest applying the ADNEX model. If the ADNEX model results were higher than the suggested cut-off, we recommend immediate referral to the oncologic unit. In the case of a negative SR or negative ADNEX model, the management of patients will depend on other factors (like symptoms or fertility desire). Larger multi-center prospective trials are needed to further evaluate the performance of these models and to establish the performance of this two-step strategies also in the hand of non-expert sonographer compared to SA of an expert sonographer.

Financial disclosure

The authors did not report any potential conflicts of interest.

CRediT authorship contribution statement

Giulia Spagnol: Conceptualization, Data curation, Writing – original draft. **Matteo Marchetti:** Data curation. **Orazio De Tommasi:** Data curation. **Amerigo Vitagliano:** Writing – review & editing. **Francesco Cavallin:** Formal analysis. **Roberto Tozzi:** Supervision, Validation, Visualization. **Carlo Saccardi:** Conceptualization, Supervision, Validation. **Marco Noventa:** Conceptualization, Writing – original draft.

Declaration of Competing Interest

The authors did not report any potential conflicts of interest.

References

[1] J. Yazbek, S.K. Raju, J. Ben-Nagi, T.K. Holland, K. Hillaby, D. Jurkovic, Effect of quality of gynaecological ultrasonography onmanagement of patients with suspected ovarian cancer: a randomised controlled trial, Lancet Oncol. 9 (2) (2008) 124–131.

[2] Y.L. Woo, M. Kyrgiou, A. Bryant, T. Everett, H.O. Dickinson, Centralisation of services for gynaecological cancers - a Cochrane systematic review, Gynecol. Oncol. 126 (2) (2012) 286–290.

[3] W. Froyman, C. Landolfo, B. De Cock, et al., Risk of complications in patients with conservatively managed ovarian tumours (IOTA5): a 2-year interim analysis of a multicentre, prospective, cohort study, Lancet Oncol. 20 (3) (2019) 448–458.

[4] J. Kaijser, T. Bourne, L. Valentin, A. Sayasneh, C. Van Holsbeke, I. Vergote, A. Testa, D. Franchi, B. Van Calster, D. Timmerman, Improving strategies for diagnosing ovarian cancer: a summary of the International Ovarian Tumor Analysis (IOTA) studies, Ultrasound Obstet. Gynecol. 41 (2013) 9–20.

[5] D. Timmerman, L. Valentin, T.H. Bourne, W.P. Collins, H. Verrelst, I. Vergote, IOTA Group, Terms, definitions, and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) Group, Ultrasound Obstet. Gynecol. 16 (2000) 500–505.

[6] I. Jacobs, D. Oram, J. Fairbanks, J. Turner, C. Frost, J.G. Grudzinskas, A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer, Br. J. Obstet. Gynaecol. 97 (10) (1990 Oct) 922–929.

[7] S. Tingulstad, B. Hagen, F.E. Skjeldestad, M. Onsrud, T. Kiserud, T. Halvorsen, K. Nustad, Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses, Br. J. Obstet. Gynaecol. 103 (8) (1996 Aug) 826–831.

[8] S. Tingulstad, B. Hagen, F.E. Skjeldestad, T. Halvorsen, K. Nustad, M. Onsrud, The risk-of-malignancy index to evaluate potential ovarian cancers in local hospitals, Obstet. Gynecol. 93 (3) (1999 Mar) 448–452.

[9] Y. Yamamoto, R. Yamada, H. Oguri, N. Maeda, T. Fukaya, Comparison of four malignancy risk indices in the preoperative evaluation of patients with pelvic masses, Eur. J. Obstet. Gynecol. Reprod. Biol. 144 (2) (2009 Jun) 163–167.

[10] TestaAC TimmermanD, T. Bourne, et al., Simple ultrasoundbased rules for the diagnosis of ovarian cancer, Ultrasound Obstet. Gynecol. 31 (2008) 681–690.

[11] D. Timmerman, B. Van Calster, A. Testa, L. Savelli, D. Fischerova, W. Froyman, L. Wynants, C. Van Holsbeke, E. Epstein, D. Franchi, J. Kaijser, A. Czekierdowski, S. Guerriero, R. Fruscio, F.P.G. Leone, A. Rossi, C. Landolfo, I. Vergote, T. Bourne, L. Valentin, Predicting the risk of malignancy in adnexal masses based on the simple rules from the international ovarian tumor analysis group, Am. J. Obstet. Gynecol. 214 (2016) 424–437.

[12] B. Van Calster, K. Van Hoorde, L. Valentin, A.C. Testa, D. Fischerova, C. Van Holsbeke, L. Savelli, D. Franchi, E. Epstein, J. Kaijser, V. Van Belle, A. Czekierdowski, S. Guerriero, R. Fruscio, C. Lanzani, F. Scala, T. Bourne, D. Timmerman, Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study, BMJ. 349 (2014) g5920.

[13] R.F. Andreotti, D. Timmerman, L.M. Strachowski, W. Froyman, B.R. Benacerraf, G.L. Bennett, T. Bourne, D.L. Brown, B.G. Coleman, M.C. Frates, S.R. Goldstein, U.M. Hamper, M.M. Horrow, M. Hernanz-Schulman, C. Reinhold, S.L. Rose, B.P. Whitcomb, W.L. Wolfman, P. Glanc, 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 Jan) 168–185.

[14] D. Timmerman, F. Planchamp, T. Bourne, C. Landolfo, A. du Bois, L. Chiva, D. Cibula, N. Concin, D. Fischerova, W. Froyman, G. Gallardo Madueño, B. Lemley, A. Loft, L. Mereu, P. Morice, D. Querleu, A.C. Testa, I. Vergote, V. Vandecaveye, G. Scambia, C. Fotopoulou, ESGO/ISUOG/IOTA/ESGE consensus statement on pre-operative diagnosis of ovarian tumors, Int. J. Gynecol. Cancer 31 (7) (2021 Jul) 961–982.

[15] C. Landolfo, T. Bourne, W. Froyman, B. Van Calster, J. Ceusters, A.C. Testa, L. Wynants, P. Sladkevicius, C. Van Holsbeke, E. Domali, R. Fruscio, E. Epstein, D. Franchi, M.J. Kudla, V. Chiappa, J.L. Alcazar, F.P.G. Leone, F. Buonomo, M.E. Coccia, S. Guerriero, N. Deo, L. Jokubkiene, L. Savelli, D. Fischerova, A. Czekierdowski, J. Kaijser, A. Coosemans, G. Scambia, I. Vergote, D. Timmerman, L. Valentin, Benign descriptors and ADNEX in two-step strategy to estimate risk of malignancy in ovarian tumors: retrospective validation in IOTA5 multicenter cohort, Ultrasound Obstet. Gynecol. 61 (2) (2023 Feb) 231–242.

[16] R.J. Kurman, cancer CIDRSL, WHO Classification of Tumours of Female Reproductive Organs, World Health Organization, 2014 307.

[17] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2023, URL https://www.R-project.org/.

[18] D.K. Armstrong, R.D. Alvarez, F.J. Backes, J.N. Bakkum-Gamez, L. Barroilhet, K. Behbakht, A. Berchuck, L.M. Chen, V.C. Chitiyo, M. Cristea, M. DeRosa, E.L. Eisenhauer, D.M. Gershenson, H.J. Gray, R. Grisham, A. Hakam, A. Jain, A. Karam, G.E. Konecny, C.A. Leath Iii, G. Leiserowitz, J. Liu, L. Martin, D. Matei, M. McHale, K. McLean, D.S. Miller, S. Percac-Lima, S.W. Remmenga, J. Schorge, D. Stewart, P.H. Thaker, R. Vargas, A.W. Hendrickson, T.L. Werner, E. Zsiros, M.A. Dwyer, L. Hang, NCCN Guidelines® Insights: Ovarian Cancer, Version 3.2022, J. Natl. Compr. Cancer Netw. 20 (9) (2022 Sep) 972–980.

[19] Royal College of Obstetricians and Gynaecologists, The Management of Ovarian Cysts in Postmenopausal Women, Green Top Guildelines, no. 34 2016.

[20] L. Valentin, Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses, Ultrasound Obstetr. Gynecol. Off. J. Int. So. Ultrasound Obstetr. Gynecol. 14 (1999) 273–283.

[21] E.M. Meys, J. Kaijser, R.F. Kruitwagen, B.F. Slangen, B. Van Calster, B. Aertgeerts, J.Y. Verbakel, D. Timmerman, T. Van Gorp, Subjective assessment versus ultrasound models to diagnose ovarian cancer: a systematic review and meta-analysis, Eur. J. Cancer 58 (2016) 17–29.

[22] B. Rahman, S.F. Meisel, L. Fraser, L. Side, S. Gessler, J. Wardle, A. Lanceley, Population-based genetic risk prediction and stratification for ovarian cancer: views from women at high risk, Familial Cancer 14 (1) (2015 Mar) 135–144.

[23] J.L. Alcázar, M.A. Pascual, B. Graupera, M. Aubá, T. Errasti, B. Olartecoechea, A. Ruiz-Zambrana, L. Hereter, S. Ajossa, S. Guerriero, External validation of IOTA simple descriptors and simple rules for classifying adnexal masses, Ultrasound Obstet. Gynecol. 48 (3) (2016 Sep) 397–402.

[24] A. Sayasneh, L. Ferrara, B. De Cock, S. Saso, M. Al-Memar, S. Johnson, J. Kaijser, J. Carvalho, R. Husicka, A. Smith, C. Stalder, M.C. Blanco, G. Ettore, B. Van Calster, D. Timmerman, T. Bourne, Evaluating the risk of ovarian cancer before surgery using the ADNEX model: a multicentre external validation study, Br. J. Cancer 115 (5) (2016 Aug 23) 542–548.

[25] J.J. Hidalgo, F. Ros, M. Aubá, T. Errasti, B. Olartecoechea, Á. Ruiz-Zambrana, J.L. Alcázar, Prospective external validation of IOTA three-step strategy for characterizing and classifying adnexal masses and retrospective assessment of alternative two-step strategy using simple-rules risk, Ultrasound Obstet. Gynecol. 53 (5) (2019 May) 693–700, https://doi.org/10.1002/uog.20163 (PMID: 30353585).

[26] L. Ameye, D. Timmerman, L. Valentin, D. Paladini, J. Zhang, C. Van Holsbeke, A.A. Lissoni, L. Savelli, J. Veldman, A.C. Testa, F. Amant, S. Van Huffel, T. Bourne, Clinically oriented three-step strategy for assessment of adnexal pathology, Ultrasound Obstet. Gynecol. 40 (5) (2012 Nov) 582–591.

[27] N. Nunes, G. Ambler, X. Foo, J. Naftalin, M. Widschwendter, D. Jurkovic, Use of IOTA simple rules for diagnosis of ovarian cancer: meta-analysis, Ultrasound Obstet. Gynecol. 44 (5) (2014 Nov) 503–514, https://doi.org/10.1002/uog.13437 (Epub 2014 Oct 13).

[28] H. Chen, L. Qian, M. Jiang, Q. Du, F. Yuan, W. Feng, Performance of IOTA ADNEX model in evaluating adnexal masses in a gynecological oncology center in China, Ultrasound Obstet Gynecol 54 (6) (2019 Dec) 815–822.

[29] B. Van Calster, K. Van Hoorde, W. Froyman, J. Kaijser, L. Wynants, C. Landolfo, C. Anthoulakis, I. Vergote, T. Bourne, D. Timmerman, Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors, Facts Views Vis Obgyn. 7 (2015) 32–41.

[30] L. Cao, Liu Y. WeiM, et al., Validation of American College of Radiology Ovarian-Adnexal Reporting and Data System Ultrasound (O-RADS US): analysis on 1054 adnexal masses, Gynecol. Oncol. 162 (1) (2021) 107–112.

[31] M.A.A. Basha, M.I. Metwally, S.A. Gamil, et al., Comparison of O-RADS, GI-RADS, and IOTA simple rules regarding malignancy rate, validity, and reliability for diagnosis of adnexal masses, Eur. Radiol. (2020).

[32] A.K. Hiett, J.D. Sonek, M. Guy, T.J. Reid, Performance of IOTA Simple Rules, Simple Rules risk assessment, ADNEX model and O-RADS in differentiating between benign and malignant adnexal lesions in North American women, Ultrasound Obstet. Gynecol. 59 (5) (2022 May) 668–676, https://doi.org/10.1002/uog.24777 (Epub 2022 Apr 8).

[33] S Timmerman, L Valentin, J Ceusters, AC Testa, C Landolfo, P Sladkevicius, C Van Holsbeke, E Domali, R Fruscio, E Epstein, D Franchi, MJ Kudla, V Chiappa, JL Alcazar, FPG Leone, F Buonomo, ME Coccia, S Guerriero, N Deo, L Jokubkiene, J Kaijser, G Scambia, R Andreotti, D Timmerman, T Bourne, B Van Calster, W Froyman, External Validation of the Ovarian-Adnexal Reporting and Data System (O-RADS) Lexicon and the International Ovarian Tumor Analysis 2-Step Strategy to Stratify Ovarian Tumors Into O-RADS Risk Groups, JAMA Oncol 9 (2) (2023) 225–233.