The Prospective External Validation of International Ovarian Tumor Analysis (IOTA) Simple Rules in the Hands of Level I and II Examiners

Prospektive externe Validierung der „International Ovarian Tumor Analysis (IOTA) Simple Rules“ durch Level I und II Untersucher

Authors A. Knafel[1], T. Banas[2], A. Nocun[2], M. Wiechec[2], R. Jach[2], A. Ludwin[3], M. Kabzinska-Turek[2], M. Pietrus[2], K. Pitynski[2]

Affiliations
[1] Dept. of Gynecology and Obstetrics, Jagiellonian University, Krakow, Poland
[2] Gynecologic Oncology, Jagiellonian University, Krakow, Poland
[3] Departement of Gynecologic Oncology, Jagiellonian University, 31–501 Krakow, Poland

Keywords
- IOTA
- simple rules
- adnexal tumors
- ultrasound
- genital/reproductive
- ovary
- pelvic organs
- ultrasound 2 D

Abstract

Objective: To externally validate the International Ovarian Tumor Analysis (IOTA) Simple Rules (SR) by examiners with different levels of sonographic experience defined by the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) and to assess the morphological ultrasound features of the adnexal tumors classified as inconclusive based on IOTA SR.

Materials and Methods: In the two-year prospective study adnexal tumors were assessed preoperatively with transvaginal ultrasound by examiners with different levels of experience (level 1-IOTA SR1, level 2-IOTA SR2). Additionally, an expert (level 3) evaluated all tumors by subjective assessment (SA). If the rules could not be applied, the tumors were considered inconclusive. The final diagnosis was based on the histopathological result of the removed mass. The diagnostic performance measures for the assessed model were sensitivity, specificity, negative (LR-) and positive (LR+) likelihood ratios, accuracy (ACC) and diagnostic odds ratio (DOR).

Results: 226 women with adnexal tumors scheduled for surgery were included in the stutdy. The prevalence of malignancy was 36.3 % in the group of all studied tumors and was 52.5 % in the inconclusive group (n = 40) (p = 0.215). Fewer tumors were classified as inconclusive by level 2 examiners compared to level 1 examiners [20 (8.8 %) vs. 40 (17.7 %); p = 0.008], resulting from the discrepancy in the evaluation of acoustic shadows and the vascularization within the tumor. For level 1 examiners a diagnostic strategy using IOTA SR1 +MA (assuming malignancy when SR inconclusive) achieved a sensitivity, specificity and DOR of 96.3 %, 81.9 %, 13.624 respectively. For level 2 examiners the diagnostic strategy for IOTA SR2 +MA achieved a sensitivity, specificity and DOR of 95.1 %, 89.6 %, 137,143, respectively. Adding SA by an expert (or level 3 examiner)

Zusammenfassung

Ziel: Externe Validierung der „International Ovarian Tumor Analysis (IOTA) Simple Rules (SR)“ durch Untersucher mit unterschiedlichem Maß an Ultraschall-Erfahrung nach Definition der European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) und die Bewertung der morphologischen Ultraschallmerkmale von Adnextumoren, die nach den IOTA SR als nicht-konklusiv definiert werden.

Material und Methoden: In der zweijährigen prospektiven Studie wurden Adnextumore präoperativ mit transvaginaler Sonografie durch Untersucher mit unterschiedlichem Maß an Erfahrung (Level-1- IOTA SR1, Level-2-IOTA SR2) beurteilt. Zusätzlich bewertete ein Experte (Level 3) alle Tumore durch subjektive Einschätzung (SA). Wenn die Regeln nicht angewandt werden konnten, wurden die Tumore als nicht-konklusiv betrachtet. Die Enddiagnose basierte auf den histopathologischen Erbebnissen in der entfernten Raumforderung. Die für das Modell beurteilten diagnostischen Leistungsparameter waren Sensitivität, Spezifität, negative (LR-) und positive (LR+) Likelihood-Ratio, Genauigkeit (ACC) and diagnostische Odds Ratio (DOR).

Ergebnisse: In die Studie wurden 226 Frauen mit Adnextumoren und geplanter Operation eingeschlossen. Die Prävalenz für Malignität in der Gruppe aller untersuchter Tumore betrug 36,3 % und innerhalb der nicht-konklusiven Gruppe (n = 40) 52,5 % (p = 0,215). Von den Untersuchern mit Level 2 wurden weniger Tumore als nicht-konklusiv beurteilt, im Vergleich zu Level-1-Untersuchern [20 (8,8 %) vs. 40 (17,7 %); p = 0,008], was durch die unterschiedlichen Bewertung der akustischen Schatten und der Vaskularisierung im Tumor bedingt war. Bei Level-1-Untersuchern erreichte die diagnostische Strategie mittels IOTA SR1 +MA (Malignität angenommen bei nicht-konklusiven SR) eine Sensitivität von 96,3 %, eine

received 12.7.2013
accepted 4.12.2014

Bibliography
DOI http://dx.doi.org/10.1055/s-0034-1398773
Published online: June 30, 2015
Ultraschall in Med © Georg Thieme Verlag KG Stuttgart · New York · ISSN 0172-4614

Correspondence
Dr. Anna Knafel
Dept.of Gynecology and Obstetrics, Jagiellonian University
Kopernika 23
31–501 Kraków
Poland
Tel.: ++ 48/12/4 24 85 60
Fax: ++ 48/12/4 24 85 84
annaknafel@yahoo.co.uk

when IOTA SR were not applicable improved the specificity of the test and achieved a DOR of 505.137 (SR1 +SA) and 293.627 (SR2 +SA). The SA by an expert proved to have the best diagnostic performance with a DOR of 5768.857, and a sensitivity and specificity of 97.6 % and 99.3 % respectively. Within the inconclusive group the most common tumors were unilocular-solid (n-13), solid (n-8) and multilocular-solid (n–10) ones. All multilocular tumors were classified as inconclusive because of their size (≥ 100 mm) and were found to be benign by pathology. Most of the inconclusive tumors with cystic content presented low-level (43.75 %) echogenicity, followed by ground-glass (34.37 %), mixed (12.5 %) and anechoic (9.4 %).

Conclusion: The study results show excellent diagnostic performance of IOTA Simple Rules followed by subjective expert assessment in inconclusive tumors irrespective of the level of experience, while subjective assessment by an expert still has the highest diagnostic odds ratio. The number of inconclusive cases seems to depend on the level of ultrasound expertise and less experienced examiners have a tendency to overestimate blood flow and a presence of acoustic shadows within the tumors. IOTA SR were not applicable either because no benign or malignant features were found or both were identified. Within inconclusive tumors the majority of cases comprise malignant masses that are either unilocular-solid, solid tumors or small multilocular-solid ones with a diameter of less than 100 mm.

Spezifität von 81,9 % und eine DOR von 13,624. Bei Level-2-Untersuchern erreichte die diagnostische Strategie mit IOTA SR2 +MA eine Sensitivität von 95,1 %, eine Spezifität von 89,6 % und eine DOR von 137,143. Die zusätzliche SA eines Experten (oder Level 3-Untersuchers) bei nicht anwendbaren IOTA SR verbesserte die Testspezifität und erreichte eine DOR von 505,137 (SR1 +SA) bzw. 293,627 (SR2 +SA). Die SA eines Experten zeigte die beste diagnostische Leistung mit einer DOR von 5768,857, einer Sensitivität von 97,6 % und einer Spezifität von 99,3 %. Innerhalb der nicht-konklusiven Gruppe sind die häufigsten Tumore unilokulär-solide (n = 13), solide (n = 8) und multilokulär-solide (n = 10). Alle multilokulären Tumore wurden aufgrund ihrer Größe (≥ 100 mm) als nicht-konklusiv klassifiziert und stellten sich in der Pathologie als gutartig heraus. Die meisten der nicht-konklusiven Tumore mit zystischem Inhalt zeigten eine geringgradige Echogenität (43,75 %), gefolgt von milchglasartiger (34,37 %), gemischter (12,5 %) und echofreier (9,4 %) Echogenität.

Schlussfolgerung: Die Studienergebnisse zeigen eine exzellente diagnostische Leistung der IOTA Simple Rules mit nachfolgender subjektiver Beurteilung des Experten bei nicht-konklusiven Tumoren unabhängig von der Expertise. Dennoch hat die subjektiven Bewertung des Experten die höchste diagnostische Odds Ratio. Die Anzahl der nicht-konklusiven Fälle scheint von der Ultraschall-Qualifikation abzuhängen und weniger erfahrene Untersucher neigen zu einer Überbewertung des Blutflusses und des Auftretens von akustischen Schatten in den Tumoren. IOTA SR waren nicht anwendbar, wenn weder benigne oder maligne Merkmale gefunden wurden oder beides entdeckt wurde. Bei nicht-konklusiven Tumoren hatte die Mehrzahl der Fälle maligne Raumforderungen, die entweder unilokulär-solide oder solide Tumore oder kleine multilokuläre Tumore mit einem Durchmesser unter 100 mm waren.

Introduction

Adnexal tumors are often diagnosed in women in daily medical practice. In order to differentiate benign from malignant lesions, many ultrasound scoring systems have been proposed so far [1 – 4]. However, none of them has been shown to be superior to the subjective assessment of grayscale and color Doppler findings, known as "pattern recognition, performed by an experienced ultrasound examiner. The accuracy of this method is as high as 89 – 96 % [5 – 7]. Unfortunately, knowledge and experience in the field of oncological ultrasound cannot be easily taught. Furthermore, access to a specialist center is often limited. These two factors underline the need to develop a method that could help less experienced sonographers to differentiate malignant from benign adnexal lesions. A set of simple ultrasound-based rules (SR) [8] defined by the International Ovarian Tumor Analysis (IOTA) Group for ovarian tumor characterization with five features of benignity and five features of malignancy proved to be a valuable tool in the preoperative assessment of adnexal masses with good results on temporal and external validation [9 – 11]. In recent studies, the performance of this model has been evaluated with respect to examiners with different levels of training and experience [12 – 14]. Unfortunately, despite the utility of IOTA SR, approximately 25 % of tumors cannot be classified according to these criteria. In these cases, referral to a specialist is recommended [9]. The question that may be raised is: what kind of tumors fall into the category of inconclusive lesions? The awareness of the ultrasound features of adnexal masses which belong in this group could aid in the discrimination of tumors by a less experienced sonographer.

The first purpose of the study was to assess the diagnostic performance of IOTA SR, according to the examiner's experience. Furthermore, we aimed to prospectively validate the performance of this method in the preoperative management of adnexal masses. Finally, the study investigated the morphological ultrasound features of adnexal tumors classified as inconclusive, based on IOTA SR.

Materials and Methods

This was a prospective study conducted in the Department of Gynecology and Oncology at the Jagiellonian University, Krakow, between January 2011 and October 2012. The study was approved by the Faculty's Research Ethics Committee. For the purpose of the study, ultrasound examiners were classified as level 1, 2 and 3, according to the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) Education and Professional Standards Committee guidelines [15]. Before the beginning of the study, a theoretical and practical half-day training was conducted for the ultrasound examiners(level 1 and 2) that participated in the study. During the session, ultrasound features adapted in IOTA simple rules were revised and discussed.

Patient and data collection

The inclusion criteria for the study were as follows: 1) age: 18 years old or older, 2) presence of an adnexal tumor requiring surgical treatment. In cases of bilateral masses, the tumor with the most complex ultrasound features was included. If the tumors presented a similar morphology, the mass with the biggest diameter or most accessible on ultrasound was included [9].

The exclusion criteria were: 1) pregnancy, 2) withdrawal of consent to participate in the study, 3) lack of histopathology result as an outcome, 3) surgery performed after more than 90 days since the initial diagnosis. Clinical data of each patient was recorded and comprised: age, parity, menopausal status, previous hysterectomy, personal and familial history of breast and ovarian cancer.

Study protocol

The first ultrasound exam was performed during the initial visit in the Outpatient Gynecologic Clinic by a level 1 examiner (MP), who recorded the result, according to IOTA SR (IOTA SR1). After establishing the initial diagnosis, all patients were admitted to the gynecological ward for further diagnosis and treatment. In the ward, the second ultrasound was performed by a level 2 examiner with the assistance of a level 3 examiner. Both physicians recorded the results separately without knowledge of each other's outcomes. Additionally, they were unaware of the results of the level 1 examiner. A level 2 (MKT) examiner evaluated the adnexal tumor with IOTA SR (IOTA SR2) and a level 3 examiner (AK) was asked to give subjective assessment (SA) of the mass as malignant or benign. The final decision regarding the type of surgical treatment was made together with the surgeons (KP, RJ, AL) and was based on the pattern recognition result (level 3 examiner), the bimanual and rectal examination and on the results of additional laboratory tests.

Ultrasound examination

During ultrasound, all patients were scanned transvaginally (TVS). In the case of large tumors, a transabdominal scan (TAS) was performed as well. Ultrasound examinations were performed by means of 2 D and static 3D using a Voluson E6 BT10 ultrasound scanner equipped with a volumetric RIC 5 – 9 MHz (transvaginal) and RAB 2 – 5 MHz (transabdominal) transducers (GE Healthcare, Zipf, Austria). In all scans, we standardized our basic 2 D settings using high frequencies with harmonics, level 5 speckle reduction and level 1 cross beam compounding. Power Doppler settings with quality set to high, wall motion filter (WMF) set to low 1 and a pulse repetition frequency (PRF) of 0.9 kHz were used.

TVS was performed according to the standards published by the IOTA group [16] and included: the origin, position, type of the lesion, size, morphological features (echogenicity, presence of solid components, papillary projections, number of locules) together with measurements and quantitative assessment of the tumor (size of the lesion in three diameters, number of locules, size of the solid components). Vascular features were examined by power Doppler imaging, subjective assessment was also performed according to the amount of blood flow detected with score 1 – 4 (none, minimal, moderate or intense, respectively).

The final ultrasound classification (for level 1/2 examiners) based on IOTA SR [8] (**Table 1**) was used to group adnexal masses into malignant or benign lesions, depending on which of the 10 ultrasound features was observed. If one or more M-features (malignant) were present in the absence of B-features (benign), the tumor was classified as malignant. If one or more B-features were present in the absence of M-features, the tumor was classified as benign. When both B-features and M-features were present or none of the features was present, the mass was considered inconclusive according to the IOTA SR criteria [8]. In these cases, the diagnosis was based on the SA conducted by an experienced examiner (level 3). For every patient included in the study, three ultrasound results were stored (IOTA SR1 and IOTA SR2 and SA). For the purpose of the study for IOTA SR, two approaches were used: the first, in which all inconclusive cases were classified as benign or malignant using SA by a level 3 examiner (SR1 +SA, SR2 +SA) and the second one, in which all inconclusive tumors were classified as malignant to decrease the number of possibly missed ovarian cancers (IOTA SR1 +MA, IOTA SR2 +MA) [12].

Table 1 IOTA SR[1] for the diagnosis of ovarian cancer. If one or more M-features apply in the absence of a B-feature, the mass is classified as malignant. If one or more B-features apply in the absence of an M-feature, the mass is classified as benign. If both M-features and B-features apply, the mass cannot be classified. If no features are present, the mass cannot be classified.

features for predicting a malignant tumor (M-features)		features for predicting a benign tumor (B- features)	
M1	irregular solid tumor	B1	unilocular
M2	presence of ascites	B2	presence of solid component where the largest solid component has a diameter < 7 mm
M3	> 4 papillary structures	B3	presence of shadows
M4	irregular multilocular-solid tumor with largest diameter ≥ 100 mm	B4	smooth multilocular tumor with largest diameter < 100 mm
M5	very strong blood flow (color score 4)	B5	no blood flow (color score 1)

[1] IOTA SR – International Ovarian Tumor Analysis Simple Rules

Surgery and pathology analysis

All patients with a confirmed adnexal mass underwent surgical treatment at our institution. The surgical procedures were chosen and performed according to the medical indication. The gold standard was the histopathological diagnosis of surgical specimens, all performed in the Department of Pathomorphology of the Jagiellonian University, following the guidelines of the World Health Organization International Classification of Ovarian Tumors [17]. For study purposes, borderline tumors were classified as malignant.

Statistical analysis

Using the Shapiro–Wilk test, the distributions of variables in the patient subgroups were analyzed. Since all variables had normal distributions, the Student's t-test was applied to compare the subgroups of patients. Parity was compared using the Chi [2] test. The clinical features of the study patients were presented as median values and standard deviation (SD) or number of cases and percentage. Subsequently, the sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) were calculated separately for each diagnostic algorithm. The positive likelihood ratio (LR+) was calculated as sensitivity/(1-specificity); the negative likelihood ratio (LR-) was calculated as (1-sensitivity)/specificity. LR+/LR- was presented as the diagnostic odds ratio (DOR). A p-value of 0.05 was accepted as statistically significant. All calculations were carried out with the use of STATISTICA data analysis software version 9.0 (TB) (www.statsoft.com).

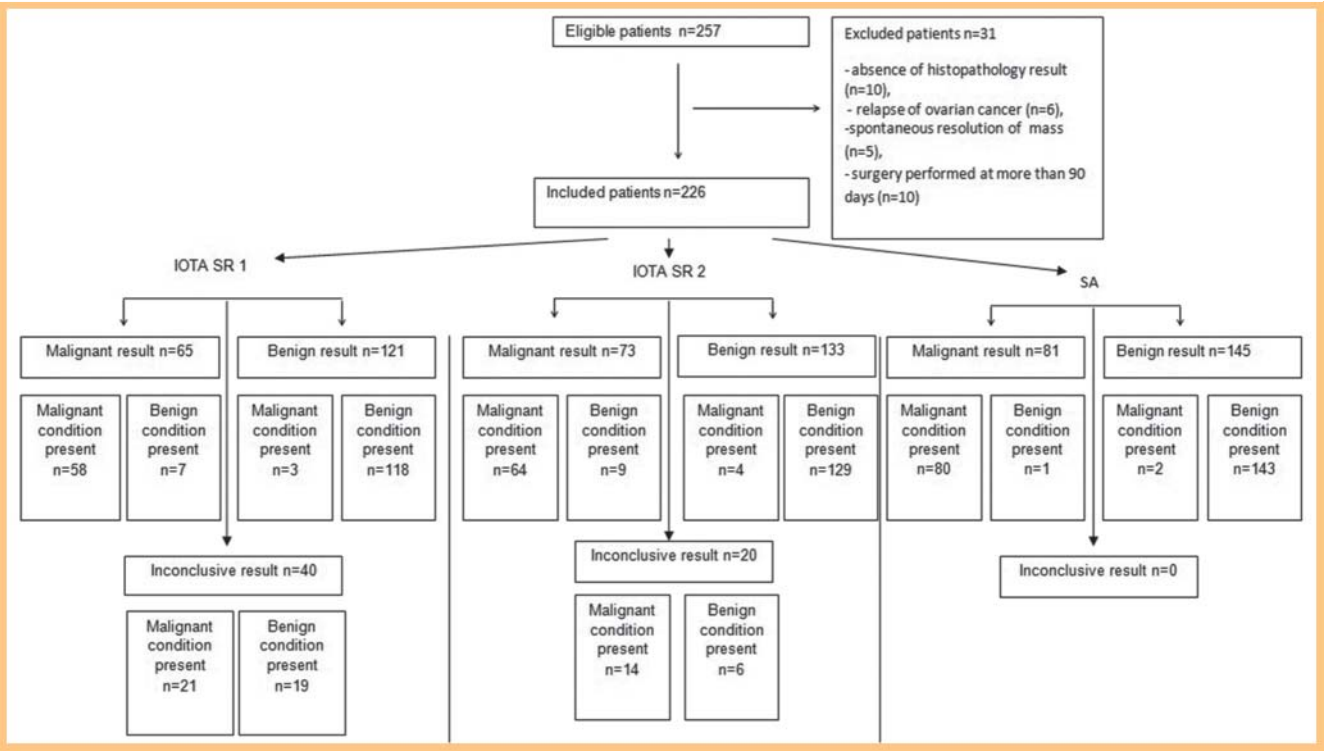

Fig. 1 Flow diagram of *IOTA SR in the hands of the examiners with different levels of experience (level 1-IOTA SR*1 and level 2-IOTA SR*2) and subjective assessment (SA) by experienced ultrasound examiners (level 3). *IOTA SR – International Ovarian Tumor Analysis Simple Rules.

Abb. 1 Flussdiagramm der *IOTA SR bei Untersuchern mit unterschiedlichem Maß an Erfahrung (Level-1-IOTA SR*1 und Level-2-IOTA SR*2) und subjektive Einschätzung (SA) von Ultraschall-Experten (Level 3). *IOTA SR – „International Ovarian Tumor Analysis Simple Rules".

The results were presented according to STARD (Standards for Reporting Diagnostic Accuracy) guidelines [18].

Results

A total of 257 women were recruited for the study. 31 patients were excluded due to: absence of a histopathology result (only cytology available) (n = 10), relapse of ovarian cancer (n = 6), spontaneous resolution of mass (n = 5), surgery performed after more than 90 days (n = 10). The final database consisted of 226 tumors (**Fig. 1**). The prevalence of malignancy was 36.3 % (82 malignant vs. 144 benign). 17 (7.5 %) tumors were bilateral. The mean age of the patients was 47 years. 143 (63.5 %) patients were premenopausal. Malignancy was found in 33 (23 %) premenopausal and in 49 (59.7 %) postmenopausal patients. Of 82 malignancies, 67 were primary epithelial ovarian cancers, 7 were borderline tumors, and 8 were metastatic tumors. The histological type and distribution of benign and malignant ovarian masses are presented in **Table 2**.

IOTA SR could be applied in 186 (82.3 %) cases by a level 1 sonographer (IOTA SR1) and in 206 (91.2 %) by a level 2 examiner (IOTA SR2). A level 2 sonographer qualified significantly less ovarian tumors as inconclusive according to IOTA SR, when compared to a level 1 examiner [20 (8.8 %) vs. 40 (17.7 %); p = 0.008]. Out of 65 tumors predicted to be malignant by a level 1 examiner, 58 (89.2 %) were malignant according to histology. For the level 2 examiner, 64 (87.7 %) tumors were malignant out of the 73 predicted to be malignant by the SR. Among the tumors for which IOTA SR yielded a conclusive result, they had a sensitivity and specificity for a level 1 examiner of 95.4 % (58/61) and 94.4 % (118/125), respectively, compared to 94.1 % (64/68) and 93.5 % (129/138), respectively, for a level 2 examiner. The IOTA SR1 missed 3 cancers and gave 7 false-positive diagnoses compared to 4 missed cancers and 9 false-positive diagnoses for IOTA SR2. The SA alone for all tumors had the highest sensitivity of 97.6 % (80/82), specificity of 99.3 % (143/144), NPV of 98.6 % and PPV of 98.8 %. If the simple rules were used as a triage test and subjective assessment of ultrasound findings was applied for inconclusive tumors, the test performance for IOTA SR1 +SA was as follows: sensitivity 96.3 % and specificity 95.1 %% compared to 95.1 % and 93.8 %, respectively, for IOTA SR2 +SA (**Table 3**). The strategy of classifying all SR inconclusive tumors as malignant (SR+MA) resulted in the same sensitivity for the examiners (IOTA SR1 +MA- 96.3 %, IOTA SR2 +MA- 95.1 %) compared to SR+SA. However, the specificity of this approach was lower with 81.9 % and 89.6 % for level 1 and 2 examiners, respectively. IOTA SR had a better performance for the tumors in premenopausal patients for the level 1 and 2 examiners irrespective of the approach chosen for the inconclusive tumors (IOTA SR+SA vs. IOTA SR+MA) (**Table 3**).

Ultrasound features of inconclusive tumors

Out of 40 inconclusive results according to IOTA SR1, 14 presented both M and B-features and the remaining 26 presented neither an M nor B-feature (**Table 4**). The evaluation of morphology according to IOTA SR in the group of inconclusive tumors was concordant in 18 cases among examiners with different levels of experience (level 1 vs. level 2). The discrepancies between the examiners resulted from the different evaluation of acoustic shadows (B3) and the blood flow within the tumor with a tendency

benign	n	%	Malignant	n	%
endometrioma	42	29.17	*ovarian carcinoma*		
mature teratoma	23	16.20	– serous	26	31.7
hydrosalpinx	3	2.11	– clear cell	6	7.32
cystadenoma/cystadenofibroma[1]	30	21.13	– endometrioid	10	12.05
fibroma/fibrothecoma	11	7.75	– mucinous	5	6.02
functional cyst (corpus luteum, hemorrhagic cyst)	10	7.04	– undifferentiated	7	8.43
Brenner tumor	1	0.70	– carcinosarcoma	4	4.82
struma ovarii	1	0.70	– mixed[2]	2	2.41
simple cyst	13	9.03	*borderline ovarian tumors*	7	8.43
parasalpingeal cyst	3	2.11	*sex-cord stromal tumors*	2	2.41
salpingitis	5	3.52	*germ cell tumors*	5	6.02
unknown (ovarian cyst with torsion)	2	1.41	*metastatic tumors*		
	144	100	– stomach	1	3.61
			– breast	1	1.20
			– colon	4	4.82
			– lymphoma	2	2.41
			total	82	100

[1] serous and mucinous
[2] endometrioid and clear cell cancer

Table 2 Histological classification and distribution of ovarian tumors.

	ACC	Sensitivity	Specificity	PPV	NPV	LR+	LR-	DOR
IOTA SR 1 +MA								
total sample	0.872	0.963	0.819	0.872	0.752	4.383	0.045	13.624
premenopausal	0.902	1.000	0.874	0.696	1.000	7.937	0.000	N/A†
postmenopausal	0.819	0.940	0.636	0.797	0.875	2.582	0.094	27.374
IOTA SR 1 + SA								
total sample	0.956	0.963	0.951	0.919	0.979	19.635	0.039	505.137
premenopausal	0.972	1.000	0.964	0.889	1.000	27.778	0.000	N/A†
postmenopausal	0.928	0.940	0.909	0.940	0.909	10.330	0.066	156.495
IOTA SR 2 +MA								
total sample	0.916	0.951	0.896	0.839	0.970	7.500	0.055	137.143
premenopausal	0.930	0.969	0.919	0.775	0.990	11.963	0.034	354.644
postmenopausal	0.892	0.940	0.818	0.892	0.887	5.165	0.073	70.414
IOTA SR 2 + SA								
total sample	0.942	0.951	0.938	0.897	0.971	15.339	0.052	293.627
premenopausal	0.958	0969	0.955	0.861	0.991	21.533	0.032	663.366
postmenopausal	0.916	0.940	0.879	0.922	0.906	7.769	0.068	113.810
SA								
total sample	0.987	0.976	0.993	0.988	0.986	139.429	0.024	5768.857
premenopausal	1.000	1.000	1.000	1.000	1.000	139.429	0.024	5768.875
postmenopausal	0.964	0.960	0.970	0.980	0.941	N/A†	0.000	N/A†

ACC-accuracy; PPV-positive predictive value; NPV-negative predictive value; IOTA SR – International Ovarian Tumor Analysis Simple Rules; SA- subjective assessment; MA- malignancy assumption; N/A – not applicable; LR+ positive likelihood ratio; LR- negative likelihood ratio; DOR – diagnostic odds ratio.

Table 3 Clinical value of ovarian tumor evaluation (sensitivity, specificity, LR+, LR- and ACC) for diagnostic models in all tumors, premenopausal group and postmenopausal group with respect to the level of experience (level 1 vs. level 2 examiner).

to overstimate these features by less experienced examiner. Within the group of inconclusive tumors, 13 were unilocular-solid, 10 were multilocular-solid, followed by solid tumors in 8 cases, 6 multilocular tumors, and 3 unilocular ones. Malignancy was found in 7 unilocular-solid tumors, 6 multilocular-solid tumors, 7 solid tumors, 0 multilocular tumors, and 1 unilocular tumor. Bilateral tumors were found in 9 cases. Most of the tumors with cystic content presented low-level echogenicity (43.75 %), followed by ground-glass (34.37 %), mixed (12.5 %) and anechoic (9.4 %). The median largest diameter was found in the group of multilocular and unilocular tumors (**Table 5**). None of the tumors presented a very strong blood flow (CS 4). The details regarding morphology are summarized in **Table 4, 5**.

Discussion

In this study, we have shown that the IOTA SR are a valuable method in the assessment of adnexal masses irrespective of the examiner's experience. As not all tumors can be evaluated by IOTA SR, we aimed to assess what kind of masses fall into this category. We found that the number and morphologic evaluation of inconclusive tumors according to IOTA SR varied depending on the level of ultrasound training but did not influence the final classification. Out of 40 tumors for which IOTA SR could not be applied after initial evaluation by a level 1 examiner, 50 % were in fact misclassified because of incorrectly scored tumor perfusion or false positivity of acoustic shadows. In the group of inconclusive cases, 18 lesions had concordant evaluation by level 1 and 2 examiners. We observed that most of the tumors within this

no.	age	IOTA SR1*	IOTA SR2*	SA**	ultrasound features				histopathology
					morphology	diameter (mm)	color score	bilateral	
1	59	M2 B3	M2	1	M	200	2	N	serous cystadenoma
2	48	M2 M4 B3	M2, M4	2	MS	90/80	2	Y	endometrioid carcinoma
3	31	M1 B3	M1 B3	2	S	81/55	2	N	undifferentiated carcinoma
4	53	M2, B1	M2 B1	2	U	113	1	N	serous carcinoma
5	82	M1 M2B5	M1 M2	2	S	30/45	3	Y	serous carcinoma
6	46	M1 B3	B3 B5	1	S	90	2	N	fibroma
7	28	M1 B3	M1	2	S	86	3	N	serous BOT
8	64	M1M2B3	M1 M2	2	S	60/40	2	Y	serous carcinoma
9	50	M1B3	M4B3	2	MS	80/90	3	Y	mixed carcinoma
10	16	M4B3	M4	1	MS	75	1	N	teratoma with necrosis
11	55	M4B3B5	B5	1	U	265	1	N	Teratoma with necrosis
12	60	M1M5B3	M1B3	2	S	88	3	Y	undifferentiated carcinoma
13	54	M4B5	M4B5	1	MS	84	2	N	serous cystadenofibroma
14	55	M1B5	M1B5	2	S	150	1	N	undifferentiated carcinoma
15	70	0	B5	1	US	45	1	N	struma ovarii
16	49	0	0	2	MS	85	2	N	clear cell carcinoma
17	61	0	0	2	MS	60	2	N	clear cell carcinoma
18	33	0	B5	1	U	150	1	N	salpingitis
19	44	0	B5	1	US	78	2	N	mucinous BOT of intestinal type
20	48	0	B5	1	M	290	1	N	mucinous cystadenoma
21	52	0	0	1	M	200	2	N	mucinous cystadenofibroma
22	60	0	B5	1	M	170	2	N	mucinous cystaednofibroma
23	67	0	0	2	MS	54/57	3	Y	serous carcinoma
24	47	0	B5	1	US	70	1	N	Brenner Tumor
25	71	0	0	2	US	70	3	N	serous carcinoma
26	64	0	0	1	MS	70	2	N	salpingitis
27	64	0	B5	1	US	87	1	N	mucinous cystadenoma
28	64	0	0	2	US	160	2	N	clear cell carcinoma
29	71	0	0	2	MS	80	2	N	endometrioid carcinoma
30	38	0	0	1	M	130	1	N	mucinous cystadenoma
31	60	0	B5	1	US	87	1	N	serous cystadenofibroma
32	29	0	0	1	US	70	1	N	endometrioma
33	24	0	M5	2	US	50	3	N	mucinous carcinoma
34	65	0	0	2	US	120	3	N	undifferentiated carcinoma
35	51	0	B5	1	US	50	1	N	teratoma
36	85	0	B5	1	M	113	1	N	mucinous cystadenoma
37	46	0	0	2	US	110	3	N	endometrioid carcinoma
38	48	0	0	2	US	43/52	2	Y	endometrioid carcinoma
39	68	0	0	1	MS	30	1	N	serous cystadenofibroma
40	62	0	M1	2	S	45/57	2	Y	serous carcinoma

0: no M and B-features present according to, *IOTA **SR, U: unilocular, S: solid, US: unilocular-solid, MS: multilocular-solid, M: multilocular, M1: irregular solid tumor, M2: presence of ascites, M4: irregular multilocular-solid tumor with largest diameter ≥ 100 mm, M5: very strong blood flow (color score 4), B1: unilocular tumor, B3: presence of acoustic shadows, B5: no blood flow (color score 1); **SA 1: benign, 2: malignant, *IOTA SR: *IOTA-International Ovarian Tumor Analysis Simple ultrasound-based Rules, **SA: subjective assessment, BOT: borderline ovarian tumor

group in which no B-rule or M-rule could be ascribed were found to be either unilocular-solid tumors, multilocular-solid with a diameter of less than 100 mm or smooth multilocular tumors with a diameter of 100 mm or more. In tumors that exhibited at least one B-rule and one M-rule, the following findings were found: irregular solid tumors (M1) with acoustic shadows (B3) or no blood flow (B5), multilocular-solid tumors with a diameter ≥ 100 mm (M4) with B3 or B5 feature present or unilocular lesions (B1) with ascites (M2). The strength of our study is that it was performed prospectively and that the examiners with different levels of experience were not familiar with each other's results. Furthermore, we evaluated IOTA SR with respect to menopausal status, in order to assess if it changes the diagnostic performance of the method in the hands of different examiners. Indeed, according to our observations, IOTA SR are a better method for the evaluation of adnexal masses in premenopausal patients than in postmenopausal ones irrespective of the examiner's level of experience.

The limitation of the study is that there was a small number of ultrasound examiners with similar experience. Therefore, interobserver variability cannot be evaluated. Furthermore, it was difficult to classify the operators' experience. We drew from EFSUMB guidelines [15] which rely on the level of practice and abilities achieved rather than on a strict number of scans performed. However, it is sometimes difficult to precisely define the boundaries between the levels of experience. In recommendations published by the Royal College of Radiologists in the UK, this problem was underlined as well [19].

In recent publications, the test performance of IOTA prediction models and rules was found to be maintained when performed

Table 5 Ultrasound features of inconclusive tumors according to IOTA SR in respect of the type of tumor.

variable	unilocular-solid	unilocular	solid	multilocular	multilocular-solid
number (%)	13 (32.5)	3 (7.5)	8 (20)	6 (15)	10 (25)
bilateral masses, n (%)	0	0	5	0	4
malignant n(%)/benign n (%)	7 (17.5)/6 (15)	1 (2.5)/2 (5)	7 (17.5)/1 (2.5)	0/6 (15)	6 (15)/4 (10)
color score					
1 (none)	6	3	1	3	2
2 (minimal)	3	0	4	3	6
3 (moderate)	4	0	3	0	2
4 (very strong)	0	0	0	0	0
median diameter (mm, median (range))	76.1 (43 – 120)	176 (113 – 265)	71.1 (30 – 150)	183 (130 – 290)	72.1 (30 – 90)
acoustic shadows (n)	0	1	6	0	3
cystic content					
anechoic	1	1	n/a	1	0
low-level	5	0	n/a	5	4
ground-glass	6	0	n/a	0	5
mixed	1	2	n/a	0	1
maximum diameter of the solid component (mm, median (range))	20 (13 – 27)	n/a	n/a	n/a	27 (17 – 49)
blood flow within solid component					
present (n)/absent (n)	6/7	n/a	5/3	n/a	6/4
papillary projections < 4 (n) /> 4 (n)	4/0	n/a	n/a	n/a	5/0

IOTA SR: International Ovarian Tumor Analysis Simple Rules.

by examiners with different training backgrounds and experience [12 – 14, 20]. In our study, a two-step strategy [12] was applied in order to test the performance of IOTA SR: the first, in which all inconclusive cases were classified as benign or malignant using SA by a level 3 examiner (SR1 +SA, SR2 +SA) and the second one, in which all inconclusive tumors were classified as malignant to decrease the number of possibly missed ovarian cancers (IOTA SR1 +MA, IOTA SR2 +MA). We found that the assumption of malignancy or subjective assessment by a level 3 examiner in inconclusive tumors resulted in a similar sensitivity for a level 1 and level 2 examiner. However, subjective assessment of inconclusive cases by a level 3 examiner improved the specificity of the presented two-step strategy (SR1 +SA or SR2 +SA), thus preventing overdiagnosis of malignancy. Sayasneh et al. [12] observed that the subjective assessment of the adnexal tumors tended to be better for medically trained examiners, who were non-experts, when compared to sonographers. It must be underlined that SA in our study was performed by a level 3 examiner, while in the study by Sayasneh et al., the examiners used their own SA as a second stage test. Due to the fact that in our group of inconclusive tumors over half comprised malignant masses, the approach with assumption of malignancy in this group seems to be reasonable, especially for sonographers, when access to an expert ultrasound examiner is limited. In the original report by Timmerman et al. [21], most of the tumors for which the IOTA SR yielded an inconclusive result comprised abscesses, fibromas and serous borderline stage I tumors. Alcazar et al. [20] found endometriomas, mucinous cystadenomas and invasive epithelial ovarian tumors within this group as well. We observed that 52.5 % of inconclusive cases were malignant tumors with a mean diameter of less than 100 mm in 76 % of cases. The benign lesions included teratomas with necrosis, salpingitis, fibromas and cystadenomas. Until the recent report of Alcazar et al. [20], the ultrasound morphology of inconclusive tumors had not been discussed in studies. The investigators observed that most of the tumors in the group in which no B-rule or M-rule could be applied were found to be either regular solid tumors with scanty or moderate blood flow or smooth multilocular tumors (> 100 mm) with scanty or moderate blood flow. To compare, in our study this group comprised unilocular-solid tumors, multilocular-solid with a diameter of less than 100 mm and smooth multilocular tumors (> 100 mm). In the group in which the lesions exhibited at least one B-rule and one M-rule according to Alcazar et al. [20], the following findings were recognized: irregular solid tumors with no flow or acoustic shadowing, tumors with more than four papillations but no blood flow. Similarly to the group of the investigators, this group of inconclusive tumors in our study comprised irregular solid tumors with acoustic shadows (B3) or no blood flow (B5). Furthermore, multilocular-solid tumors with a diameter of > 100 mm with a B3 or B5 feature were present and unilocular lesions with ascites were found in this group as well.

If the size of the tumor is taken into account, multilocular and multilocular-solid tumors can be classified as either benign (B4) or malignant (M4) by IOTA SR and the cut-off diameter is 100 mm. In our study, all inconclusive multilocular tumors measured over 100 mm and were found to be cystadenomas/cystadenofibromas. On the other hand, in multilocular-solid masses, the mean diameter was too small to classify them by IOTA rule M4, despite the subjective features of malignancy. These findings underline the fact that tumor size inevitably influences the applicability and the diagnostic performance of IOTA SR. The subject of tumor size was raised in a study by di Legge et al. [22], who found that the application of IOTA simple rules was different in tumors of variable size. IOTA SR sensitivity with regard to malignancy was the lowest in tumors of less than 4 cm, while the specificity was the lowest in tumors ≥ 10 cm [22].

The review of the inconclusive results showed that, in a few cases, there was a discrepancy in the evaluation of acoustic shadows.

This feature, known as B3 in IOTA SR, is suggestive of a benign nature of a tumor. However, it was inappropriately assigned to the initial ultrasound (level 1 examiner) in malignant solid masses. Acoustic shadows that are the regions of low signal intensity after boundaries with very high acoustic impedance differences appear on ultrasound as an echo-free area located behind the echogenic structure. They are most clearly observed in caliceal stones or gallstones. However, in ovarian tumors, this feature is usually more subtle and is common in benign lesions such as fibromas or teratomas [23, 24]. An important source of mistakes in the evaluation of acoustic shadows may also stem from motion artifacts, when extended view images in large tumors are evaluated. The further mistake, made by level 1 examiners in the assessment of inconclusive tumors, was the classification of the color score. In most cases, there was an overestimation of the blood flow. Due to the subjective nature of the assessment of acoustic shadows and the vascularization of a tumor, more impact should be made when training programs are considered as these features were the most common sources of mistakes in our analysis. Ultrasound training programs, such as the one recently developed by Alcazar et al. [25] with the use of 3D volumes, achieved a high diagnostic performance after analysis of 200 cases of different ovarian tumors.

Conclusion

According to our results, IOTA SR are valuable in the initial assessment of ovarian masses and maintain a high diagnostic performance in conclusive tumors. The number of inconclusive cases seems to depend on the level of experience. In these tumors, the subjective assessment by an experienced level 3 examiner is a preferred strategy.

IOTA SR were not applicable either because no B and M-rules were found or both were identified. The majority of inconclusive tumors comprise malignant masses that are either unilocular-solid, solid tumors or small multilocular-solid ones with a diameter of less than 100 mm.

References

1 *Sassone AM, Timor-Tritsch IE, Artner A et al.* Transvaginal sonographic characterization of ovarian disease: evaluation of a new scoring system to predict ovarian malignancy. Obstet Gynecol 1991; 78: 70 – 76
2 *DePriest PD, Varner E, Powell J et al.* The efficacy of a sonographic morphology index in identifying ovarian cancer: a multi-institutional investigation. Gynecol Oncol 1994; 55: 174 – 178
3 *Jacobs I, Oram D, Fairbanks J et al.* A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. Br J Obstet Gynaecol 1990; 97: 922 – 929
4 *Ferrazzi E, Zanetta G, Dordoni D et al.* Transvaginal ultrasonographic characterization of ovarian masses: comparison of five scoring systems in a multicenter study. Ultrasound Obstet Gynecol 1997; 10: 192 – 197
5 *Timmerman D, Schwärzler P, Collins WP et al.* Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. Ultrasound Obstet Gynecol 1999; 13: 11 – 16
6 *Valentin L.* Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses. Ultrasound Obstet Gynecol 1999; 14: 273 – 283
7 *Van Calster B, Timmerman D, Bourne T et al.* Discrimination between benign and malignant adnexal masses by specialist ultrasound examination versus serum CA-125. J Natl Cancer Inst 2007; 99: 1706 – 1714
8 *Timmerman D, Testa AC, Bourne T et al.* Simple ultrasound-based rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol 2008; 31: 681 – 690
9 *Timmerman D, Ameye L, Fischerova D et al.* Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: prospective validation by IOTA group. BMJ 2010; 341: c6839
10 *Kaijser J, Sayasneh A, Van Hoorde K et al.* Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: a systematic review and meta-analysis. Hum Reprod Update 2014; 20: 449 – 462
11 *Fathallah K, Huchon C, Bats AS et al.* External validation of simple ultrasound rules of Timmerman on 122 ovarian tumors. Gynecol Obstet Fertil 2011; 39: 477 – 481
12 *Sayasneh A, Wynants L, Preisler J et al.* Multicentre external validation of IOTA prediction models and RMI by operators with varied training. Br J Cancer 2013; 108: 2448 – 2454
13 *Sayasneh A, Kaijser J, Preisler J et al.* A multicenter prospective external validation of the diagnostic performance of IOTA simple descriptors and rules to characterize ovarian masses. Gynecol Oncol 2013; 130: 140 – 146
14 *Hartman CA, Juliato CR, Sarian LO et al.* Ultrasound criteria and CA 125 as predictive variables of ovarian cancer in women with adnexal tumors. Ultrasound Obstet Gynecol 2012; 40: 360 – 366
15 European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB). Minimum training requirements for the practice of medical ultrasound in Europe. www.efsumb.org/guidelines/2009-04-14apx1.pdf
16 *Timmerman D, Valentin L, Bourne TH. . International Ovarian Tumor Analysis (IOTA) Group.. et al.* 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 2000; 16: 500 – 517
17 *Kurman RJ, Carcangiu ML, Herrington CS et al.* WHO Classification of Tumours, IARC WHO Classification of Tumours. WHO Press; 2014; No 6. Edition 4
18 *Bossuyt PM, Reitsma JB, Bruns DE et al.* The Standards for Reporting of Diagnostic Accuracy Group. Standards for Reporting of Diagnostic Accuracy Group. Croat Med J 2003; 44: 639 – 650
19 RCOG Green-top Guideline No. 62. RCOG/BSGE Joint Guideline 2011. Royal College of Obstetricians and Gynaecologists. Management of Suspected Ovarian Masses in Premenopausal Women
20 *Alcázar JL, Pascual MÁ, Olartecoechea B et al.* IOTA simple rules for discriminating between benign and malignant adnexal masses: prospective external validation. Ultrasound Obstet Gynecol 2013; 42: 467 – 471
21 *Timmerman D, Van Calster B, Testa AC et al.* Ovarian cancer prediction in adnexal masses using ultrasound-based logistic regression models: a temporal and external validation study by the IOTA group. Ultrasound Obstet Gynecol 2010; 36: 226 – 234
22 *Di Legge A, Testa AC, Ameye L et al.* Lesion size affects diagnostic performance of IOTA logistic regression models, IOTA simple rules and risk of malignancy index in discriminating between benign and malignant adnexal masses. Ultrasound Obstet Gynecol 2012; 40: 345 – 354
23 *Paladini D, Testa A, Van Holsbeke C et al.* Imaging in gynecological disease (5): clinical and ultrasound characteristics in fibroma and fibrothecoma of the ovary. Ultrasound Obstet Gynecol 2009; 34: 188 – 195
24 *Wang S, Johnson S.* Predition of benignity of solid adnexal masses. Arch Gynecol Obstet 2012; 285: 721 – 726
25 *Alcázar JL, Díaz L, Flórez P.* Intensive training program for ultrasound diagnosis of adnexal masses: protocol and preliminary results. Ultrasound Obstet Gyneco 2013; 42: 218 – 223