

The use of sonographic subjective tumor assessment, IOTA logistic regression model 1, IOTA Simple Rules and GI-RADS system in the preoperative prediction of malignancy in women with adnexal masses

Jarosław Koneczny¹, Artur Czekierdowski², Marek Florczak¹, Paweł Poziemski¹,
Norbert Stachowicz³, Dariusz Borowski⁴

¹Department of Obstetrics and Gynecology, County Hospital in Minsk Mazowiecki, Poland

²Department of Gynecologic Oncology and Gynecology of the Medical University of Lublin, Poland

³Department of Epidemiology and Clinical Research Methodology, Medical University of Lublin, Poland

⁴Department of Obstetrics and Gynecology, Collegium Medicum, Nicolaus Copernicus University in Torun, Poland

ABSTRACT

Background: Sonography based methods with various tumor markers are currently used to discriminate the type of adnexal masses.

Objective: To compare the predictive value of selected sonography-based models along with subjective assessment in ovarian cancer prediction.

Material and methods: We analyzed data of 271 women operated because of adnexal masses. All masses were verified by histological examination. Preoperative sonography was performed in all patients and various predictive models including IOTA group logistic regression model LR1 (LR1), IOTA simple ultrasound-based rules by IOTA (SR), GI-RADS and risk of malignancy index (RMI3) were used. ROC curves were constructed and respective AUC's with 95% CI's were compared.

Results: Of 271 masses 78 proved to be malignant including 6 borderline tumors. LR1 had sensitivity of 91.0%, specificity of 91.2%, AUC = 0.95 (95% CI: 0.92–0.98). Sensitivity for GI-RADS for 271 patients was 88.5% with specificity of 85% and AUC = 0.91 (95% CI: 0.88–0.95). Subjective assessment yielded sensitivity and specificity of 85.9% and 96.9%, respectively with AUC = 0.97 (95% CI: 0.94–0.99). SR were applicable in 236 masses and had sensitivity of 90.6% with specificity of 95.3% and AUC = 0.93 (95% CI 0.89–0.97). RMI3 was calculated only in 104 women who had CA125 available and had sensitivity of 55.3%, specificity of 94% and AUC = 0.85 (95% CI: 0.77–0.93).

Conclusions: Although subjective assessment by the ultrasound expert remains the best current method of adnexal tumors preoperative discrimination, the simplicity and high predictive value favor the IOTA SR method, and when not applicable, the IOTA LR1 or GI-RADS models to be primarily and effectively used.

Key words: ovarian cancer prediction, sonography, IOTA group, Simple Rules, LR1 model, GI-RADS model, RMI model

Ginekologia Polska 2017; 88, 12: 647–653

INTRODUCTION

In the past two decades ultrasound examination of the female pelvis has gained a vast acceptance by both gynaecologists and perinatologists. Currently, it is the most common and non-invasive method for the uterus and the adnexa imaging and plays a very important role in the ini-

tial pelvic masses evaluation [1]. An accurate prediction of an adnexal mass type enables the choice of an appropriate surgical method with less aggressive procedures like minilaparotomy or laparoscopy indicated in presumably benign masses and immediate referral to gynecological oncologist in tumors suspected for malignancy [2]. These

Corresponding author:

Jarosław Koneczny
Department of Obstetrics and Gynecology
County Hospital in Minsk Mazowiecki, Poland
e-mail: konecznyjaroslaw@gmail.com

predictions may save patients life in cases of ovarian cancers and save costs if extensive surgery is not needed [3]. If the mass seems to be malignant then timed and appropriate treatment by gynecologic oncologist is mandatory and is related with better long-term prognosis [4–6]. Currently, subjective opinion of an ultrasound expert is regarded as the most effective method to discriminate adnexal masses before surgery [7, 8].

The advantage of various mathematical models is related to their relative simplicity and a possibility to be used by less experienced sonographers who need to triage women with adnexal masses [10]. To date, numerous scoring systems for preoperative masses discrimination have been developed. Most of these models were developed in a relatively small groups of women and later were never prospectively validated in the new populations of patients with adnexal tumors. The goal of a good prognostic model is to minimize false negative results and maximize true negative results, and both objectives are difficult to obtain at the same time. For instance, a very popular ROMA algorithm used clinically since 2011 does not appear to be better than ultrasound examination. In fact, experienced gynecologic sonography specialists can distinguish most of ovarian tumors before surgery with a significantly higher accuracy than ROMA test [10].

In 2000 IOTA group has released a consensus statement how to describe and score adnexal masses [11]. The first international and multicenter results of validated new prognostic models in women with adnexal masses were presented more than decade ago by the IOTA (International Ovarian Tumor Analysis) Group collaboration [12]. These discoveries in adnexal masses sonography had an impact on clinical practice but would not be possible before the standardization of nomenclature and ultrasound gynecological examination. Currently, the most commonly used IOTA group models include logistic regression models LR1 and LR2 and Simple Rules (SR) [13]. This latter method (SR) has been approved and published in 2015 recommendations released by the Polish Gynecological Society [14]. IOTA group models are not the only ones currently used in preoperative adnexal masses assessment. Among others, the Risk of Malignancy Index (RMI) which incorporates ultrasound score, menopausal status and serum CA125 levels was vastly used in Great Britain since its release in 1990 with further modifications [15–17]. Another, however, less popular ultrasound-based scoring system of adnexal masses is called Gynecologic-Imaging and Reporting Data System (GI-RADS). This method was first introduced by Amor et al. in 2009 [18]. These authors attempted to standardize definitions and nomenclature for all important ovarian tumor features seen during sonographic examination. The GI-RADS method was prospectively validated in a multicenter study in 2011 [19].

The aim of the current study was to compare the predictive value of selected sonography-based models along with subjective assessment in the preoperative characterization of ovarian masses.

MATERIAL AND METHODS

The study included 271 women with adnexal masses who had been preoperatively examined with transvaginal ultrasound with the use of a standardized gynecological examination method. In Lublin 101 (37.2%) cases were collected and the remaining 170 women with adnexal masses were operated in Minsk Mazowiecki county hospital. For the purpose of this study selected clinical data were recorded and included: menopausal status, age, serum CA-125 level, previous ovarian cancer history, current hormonal therapy and the presence of pain during examination. Women after hysterectomy who were over 50 years old were considered postmenopausal. Exclusion criteria were as follows: pregnancy diagnosed in woman with an adnexal tumor, failure to consent for transvaginal ultrasonography, cytology not histology as an outcome and surgical procedure that was performed more than 120 days of the date of the ultrasound examination. The predetermined cut-off of 120 days was used in order to ensure that any morphological features of a mass scored with preoperative sonography would not have had time to change substantially prior to surgery as originally set criterion for the IOTA group studies [7, 10, 12, 13]. The final outcome was the histological diagnosis of surgically removed masses. Histological results were classified according to criteria recommended by FIGO guidelines. Borderline tumors of the ovary were considered malignant for the purpose of this study. Due to retrospective nature of this study formal ethical approval was not required. Nevertheless, the study protocol was approved by the local Bioethical Committee of both centers.

Ultrasound tumor scans were performed with high-end ultrasound equipment, GE E8 EXPERT system in Lublin and GE S8 system in Minsk Mazowiecki. Transvaginal probes of these systems used ultrasound beam frequencies between 5.0 and 9.0 MHz and abdominal probes used frequencies between 3.5 and 5.0 MHz. Adnexal masses were examined with transvaginal gray-scale followed by color and “power” Doppler ultrasound examination using a standardized technique. Pulse repetition frequency (PRF) in “power” Doppler mode was set to 0.3 kHz or 0.6 kHz. During the examination of tumor vascularity, color gain was gradually increased until flash artifact appeared, then it was decreased to a level where slow flow in blood vessels was still visible. If the tumor was too large to be assessed with transvaginal probe, the transabdominal probes were also used.

All scans in Minsk Mazowiecki hospital were performed by a single gynecologist (JK) who was trained in IOTA terms and sonographic tumor criteria. This examiner at the time

Table 1. IOTA Simple Rules assessment

| Rules for predicting malignant tumor (M-rules) | Rules for predicting benign tumor (B-rules) |
|--|---|
| M1 Irregular solid tumor | B1 Unilocular cyst |
| M2 Presence of ascites | B2 Presence of solid components where the largest solid component has a largest diameter < 7 mm |
| M3 At least four papillary structures | B3 Presence of acoustic shadows |
| M4 Irregular multilocular-solid tumor with the largest diameter > 100 mm | B4 Smooth multilocular tumor with the largest diameter < 100 mm |
| M5 Very strong blood flow (color score 4) | B5 No blood flow (color score 1) |

of the study performed about 500 gynecological scans per year which is in concordance with level II examiner according to the Education and Practical Standards Committee and the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines [20]. In Lublin all scans were assessed by the level III examiner (AC) who is also an IOTA expert and one of principal investigators of IOTA studies.

Tumor sonographic morphology assessment

For the purpose of this analysis, in each women the mass with the most complex ultrasound morphology was used in case of bilateral adnexal masses. If both masses had similar sonographic features the largest tumor or the one most easily accessible with ultrasound was included. The masses were described using the terms and definitions published by the IOTA group [13]. Papillary projections were defined as projections of solid tissue into the cystic cavity arising from the cyst wall or from a septum with a height greater than or equal to 3 mm. We have measured the largest solid component other than a papillary projection. These included solid components that apparently were not protruding into the cystic mass cavity. In accordance with the IOTA consensus statement if a papillary projection was the largest solid component of a mass, the papillary projection was recorded and measured both as a papillary projection and as the largest solid component [13].

The presence of ascites and fluid in the pouch of Douglas was noted. Intraabdominal metastases were scored to be present if any extrauterine mass inconsistent with normal physiology was found in the pelvis and/or abdomen. In particular these masses had to be visible as small tumors on peritoneal surfaces when ascites was present or as separate, vascularized solid masses without ascites seen on transabdominal or transvaginal imaging.

Subjective assessment

All tumors were assessed using pattern recognition [7] with the aim to identify typical sonographic adnexal tumor patterns. The vascularization of the tumors on color Doppler was described using the IOTA color score method: no detec-

table blood flow (color score = 1), minimal blood flow (color score = 2), moderate blood flow (color score = 3) or abundant blood flow (color score = 4). The specific diagnosis of tumor histological type as suggested by the original ultrasound examiner in the first sonography report was also recorded.

IOTA Simple Rules method

Simple ultrasound based rules were assessed according to the original IOTA protocol published in 2008 [14]. The selected criteria used in this method are presented in Table 1. According to the original IOTA method, an adnexal mass was classified as malignant if at least one M-feature and no B-features were present and vice versa. When no B or M-features were present or if both B and M-features were present, then SR were considered inconclusive (uncertain), the mass was unclassifiable with this method and another second stage diagnostic method was required for these residual tumors.

RMI — Risk of Malignancy Index assessment

The third modification of RMI, also called "RMI3" was used in our analysis. According to the modified original RMI method [16] this score was calculated as below:

$$\text{RMI3}_{\text{score}} = U \times M \times \text{CA125};$$

where "U" are ultrasound features suggestive of malignancy: the presence of a multilocular cystic lesion; solid areas; bilateral lesions; ascites; intra-abdominal metastases. Zero or one point is given for each feature. If a total score is 0 or 1 point, it yields $U = 1$, a score of 2 points or more yields $U = 3$. In premenopausal women $M = 1$ and in postmenopausal patients $M = 3$. CA-125 level is used directly in this equation with the units/mL measured in blood serum. For the purpose of this study we used the "RMI3" shortcut in tables and figure's descriptions. The cut-off for malignancy is set at $\text{RMI3} = 200$ [16].

IOTA logistic regression LR1 model

The IOTA LR1 was used as originally described by Timmerman et al. [12]. This logistic regression model consists of

12 variables: (1) age of the patient (in years); (2) the presence of ascites (yes = 1, no = 0); (3) the presence of blood flow within a papillary projection (yes = 1, no = 0); (4) largest diameter of the solid component (expressed in mm but with no increase above 50 mm); (5) irregular internal cyst walls (yes = 1, no = 0); (6) the presence of acoustic shadows (yes = 1, no = 0); (7) personal history of ovarian cancer (yes = 1, no = 0); (8) current hormonal therapy (yes = 1, no = 0); (9) largest diameter of the lesion (mm); (10) the presence of pain during the examination (yes = 1, no = 0); (11) the presence of a purely solid tumor (yes = 1, no = 0); and (12) the color score (1, 2, 3 or 4).

The LR1 model's estimated probability of malignancy for an adnexal tumor equals $1/(1 + e^{-z})$, where $z = -6.7468 + 0.0326(1) + 1.5513(2) + 1.1737(3) + 0.0496(4) + 1.1421(5) - 2.3550(6) + 1.5985(7) - 0.9983(8) + 0.00841(9) - 0.8577(10) + 0.9281(11) + 0.4916(12)$, and "e" is the mathematical constant and the base value of natural logarithms. The cut-off risk for malignant masses was established as originally published at 0.1 which equals to 10% risk of malignancy [12].

GI-RADS model

All masses were scored according to the GI-RADS prognostic model that was first published by Amor et al. [18] The GI-RADS classification system is explained in Table 2. A tumor was regarded as probably malignant if it was graded 4 or 5 points.

Results were presented as absolute frequency (percentage) for nominal variables and as median (range) for continuous variables. For subjective assessment (SA), as well as for RMI3, GI-RADS, IOTA LR1 and IOTA SR Areas under the Receiver Operating Characteristic (ROC) curves (AUCs) with 95% confidence intervals (CI) were calculated to compare the discriminatory performance of the studied models. In addition, sensitivity and specificity as well as the positive and negative likelihood ratios (LR+ and LR-) were calculated

for the cutoff points proposed in the original articles. For different methods efficacy comparison receiver operating characteristic (ROC) curves were constructed and corresponding AUCs were calculated. The ROC curve displays the relationship between the test true positive ratio and the false positive ratio for a range of cutoff values. The AUC comparison was used to test various diagnostic methods ability to discriminate benign and malignant masses.

All statistical analyses were performed with the use of Stata Statistical Software, Release 11 (Stata Corp. College Station, TX, USA).

RESULTS

Of 271 masses 78 proved to be malignant including 6 borderline tumors. Table 3 presents sensitivities and specificities with related LR+ and LR- values for all used predictive methods and calculated AUC's with standard errors and 95% confidence intervals for all used prognostic methods. Preoperative serum CA-125 concentrations were available in 104 women and therefore RMI3 values could be calculated only for this group of patients. Figure 1 shows AUC's calculated for ROC curves for all studied methods. The highest predictive values were found for subjective expert assessment and for IOTA group logistic regression model 1. The LR1 model had sensitivity of 91.0%, specificity of 91.2% and AUC = 0.95 (95% CI: 0.92–0.98). Subjective assessment yielded sensitivity and specificity of 85.9% and 96.9%, respectively with AUC = 0.97 (95% CI: 0.94–0.99). For GI-RADS model sensitivity was 8.5% with specificity of 85% and AUC = 0.91 (95% CI: 0.88–0.95). Simple rules method was applicable in 236 masses (87%) and had sensitivity of 90.6% with specificity of 95.3% and AUC = 0.93 (95% CI 0.89–0.97). RMI3 had sensitivity of 55.3% with specificity of 94% and AUC = 0.85 (95% CI: 0.77–0.93). Predictive values of the studied prognostic tests for an expert (level III) and non-expert (level II) examiners are presented in Table 4. In a subgroup of premenopausal patients level II examiner

Table 2. Criteria used for GI-RADS adnexal tumors classification [18]

| GI-RADS grade | Diagnosis | Est. prob. malignancy | Detail |
|---------------|-------------------------|-----------------------|--|
| 1 | Definitive benign | 0% | Normal ovaries identified and no adnexal mass seen |
| 2 | Very probably benign | < 1% | Adnexal lesions thought to be of functional origin, e.g. follicles, corpora lutea, hemorrhagic cysts |
| 3 | Probably benign | 1–4% | Neoplastic adnexal lesions thought to be benign, such as endometrioma, teratoma, simple cyst, hydrosalpinx, paraovarian cyst, peritoneal pseudocyst, pedunculated myoma, or findings suggestive of pelvic inflammatory disease |
| 4 | Probably malignant | 5–20% | Any adnexal lesion not included in GI-RADS 1–3 and with one or two findings suggestive of malignancy* |
| 5 | Very probably malignant | > 20% | Adnexal masses with three or more findings suggestive of malignancy* |

*Thick papillary projections, thick septations, solid areas and/or ascites, defined according to IOTA criteria, and vascularization within solid areas, papillary projections or central area of a solid tumor on color or power Doppler assessment. "Est. prob." — estimated probability

Table 3. Predictive methods AUCs with 95% CIs in the differentiation between benign and malignant adnexal tumors and sensitivity, specificity with positive and negative likelihood ratios calculation in 271 women with adnexal tumors

| Method | Patients | AUC* | Std. Err. | AUC's 95% CI | | Sens. | Spec. | LR+ | LR- |
|-----------------------|----------|------|-----------|--------------|------|--------|--------|-------|------|
| Subjective assessment | 271 | 0.96 | 0.01 | 0.94 | 0.98 | 85.90% | 96.90% | 27.63 | 0.14 |
| IOTA LR1 | 271 | 0.95 | 0.013 | 0.92 | 0.97 | 91.00% | 91.20% | 10.34 | 0.09 |
| GI-RADS | 271 | 0.91 | 0.018 | 0.87 | 0.95 | 88.40% | 84.90% | 5.88 | 0.13 |
| Simple Rules | 236 | 0.92 | 0.02 | 0.89 | 0.96 | 90.60% | 95.30% | 19.48 | 0.09 |
| RMI3 | 104 | 0.84 | 0.039 | 0.77 | 0.92 | 55.20% | 93.90% | 9.11 | 0.47 |

*AUC — area under receiver-operator characteristics curve

Abbreviations: LR1 — Logistic Regression model 1; GI-RADS — Gynecologic Imaging Data and Reporting System; SR — Simple Rules have three levels (benign, inconclusive and malignant) and is represented by the ROC curve with two points; RMI3 — Risk of Malignancy Index 3; Sens. — sensitivity, Spec. — specificity; LR+ — positive likelihood ratio; LR- — negative likelihood ratio

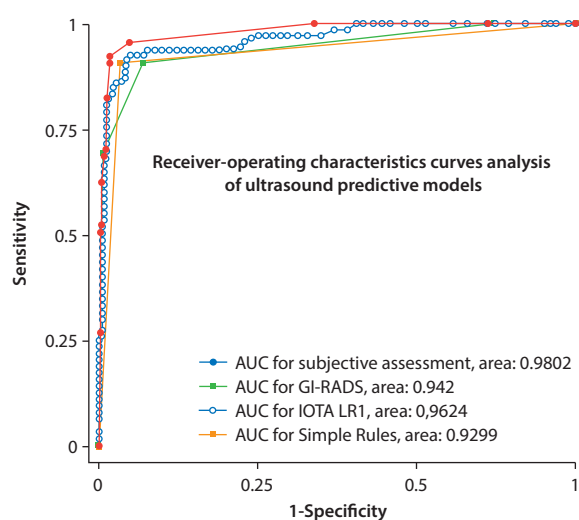


Figure 1. Comparison of receiver characteristic curves AUC's for SA, IOTA LR1, IOTA SR and GI-RADS predictive methods. LR1 — logistic regression model 1; GI-RADS — Gynecologic Data and Reporting System; SR — Simple Rules have three levels (benign, inconclusive and malignant) and is represented by a ROC curve with two points

prediction with subjective assessment yielded sensitivity 59.1% and specificity of 98%, which was much lower than for any of the ultrasound based models. The data indicate that subjective assessment predictive value was almost matched by the LR1 model while GI-RADS with SR methods had also very high predictive values in the studied group of women with adnexal masses.

The use of GI-RADS method produced relatively high number of false positive results with 29 tumors classified by level II examiner as malignant when final histology proved to be benign. False positive results were mostly found in premenopausal women. The reason of such results were probably related to the difficulties in classification of tumors with the GI-RADS score equal to 4. Similar problems and results were obtained by Amor et al. who have reported nearly 80% of false positive results if the tumors were scored 4 points [18, 19]. The lowest predictive performance was related to RMI3 model, where 17 from 38 cancers (44.7%) were incorrectly classified as being benign.

Table 4. Sensitivity and specificity of various prognostic methods as assessed by level II and level III examiners

| Examiner — level III | Sens. | Spec. | ACC | PPV | NPV |
|-----------------------|--------|--------|--------|--------|--------|
| Subjective assessment | 96.40% | 93.30% | 95.00% | 94.74% | 95.45% |
| IOTA LR1 | 94.60% | 82.20% | 89.10% | 86.90% | 92.50% |
| GI-RADS | 94.60% | 75.50% | 86.10% | 82.80% | 91.90% |
| Simple Rules | 94.10% | 94.10% | 94.10% | 96.00% | 91.40% |
| RMI3 | 62.50% | 86.60% | 71.70% | 88.20% | 59.10% |
| Examiner — level II | Sens. | Spec. | ACC | PPV | NPV |
| Subjective assessment | 59.10% | 97.90% | 92.90% | 81.20% | 94.10% |
| IOTA LR1 | 81.80% | 93.90% | 92.30% | 66.60% | 97.20% |
| GI-RADS | 72.70% | 87.80% | 85.80% | 47.00% | 95.60% |
| Simple Rules | 76.90% | 95.60% | 94.00% | 62.50% | 97.70% |
| RMI3 | 42.80% | 96.10% | 84.60% | 75.00% | 85.90% |

Abbreviations: LR1 — logistic regression model 1; GI-RADS — Gynecologic Imaging Data and Reporting System; SR — Simple Rules have three levels (benign, inconclusive and malignant) and is represented by the ROC curve with two points; RMI3 — Risk of Malignancy Index 3; Sens. — sensitivity, Spec. — specificity; ACC — accuracy; PPV — positive predictive value; NPV — negative predictive value

DISCUSSION

Accurate prognosis of the type of adnexal mass detected during ultrasound evaluation remains the most important issue in ovarian tumors classification. Prognostic ultrasound models should precisely indicate tumor type because incorrect classification may lead to an underestimation of the risk of ovarian mass malignancy. Delayed referral of a woman with suspected adnexal tumor to gynecologic oncology specialist is related to several factors, and the lack of widely-accepted and effective preoperative risk assessment strategies remains one of the most important issues [21]. In cases of false positive results unnecessary surgery is also highly probable. Unfortunately, experienced sonographic examiners are not always available and to mimic their capabilities and skills multiple ultrasound-based prognostic models have been constructed in order to help non-expert operators to more efficiently discriminate between benign and malignant ovarian tumors.

Our results indicate that IOTA Risk LR1 and GI-RADS models as well as the IOTA SR method can appropriately predict the presence of ovarian malignancy in a vast majority of adnexal masses in women who are candidates for surgical procedures. We have also confirmed that an expert could correctly classify most of the malignant ovarian tumors. Less experienced examiners should use prognostic models, preferably suggested by the IOTA group as these methods are relatively easy to learn and they retain high predictive values even in hands of level II ultrasound operators [22, 23]. In our group of patients SR method could be applied to classify 87% (236/271) of tumors. This was a higher rate than originally reported 75% of the adnexal masses which could be classified with this method [13], probably due to a relatively high percentage of malignant cases. Subjective assessment by level II examiner in our study was good for the correct preoperative classification of simple cysts, most endometriomas and mature teratomas, as well as for the majority of fibromas, tubo-ovarian abscesses and advanced FIGO stage ovarian cancers.

We have demonstrated that the IOTA group LR1 prognostic model and GI-RADS performed well when used both by an experienced and less experienced operators of ultrasound systems. Our study is another example of the external validation of the prognostic ultrasound-based models currently used in clinical practice. The results indicate that even without a special expertise less experienced sonographers are also able to correctly classify the vast majority of adnexal masses if they use well-established criteria of benign and malignant tumors. The strengths of this study are related to its multicenter design with a relatively large groups of patients diagnosed and operated in both oncologic and non-oncologic hospitals and to the use of multiple predictive models as well as the same, standardized

IOTA group terminology of adnexal masses by both expert and non-expert ultrasound examiners. Our study has also limitations which include its retrospective design and sub-optimal predictive models efficacy comparisons related to the lack of CA125 data for all operated women. Because of this missing information we were unable to calculate the RMI3 values in a substantial proportion of studied patients.

To the best of our knowledge, this is the first study in Poland that attempted to externally validate the IOTA group LR1 model and compared the results with subjective tumor assessment and other commonly used prognostic models that included RMI, SR and GI-RADS. Two groups of patients were examined preoperatively in different centers with a mixed populations of both low and high ovarian malignancies prevalence. Examiners had two levels of experience and training in ultrasound characteristics of adnexal masses. Sonographer's medium or low experience in adnexal pathology differentiation is probably the most commonly encountered situation in our country, where majority of primary ultrasound scans are performed in an outpatient setting by radiologists or gynecologists without specialized oncological training i.e. level I or level II [24]. Although rare, some of the complex adnexal masses are difficult to classify even by the experts in gynecological and oncological ultrasound who could not predict the correct diagnosis in approximately 8% of tumors [25]. It is important to note that as many as 60% of these masses turned out to be malignant following their surgical removal. This observation suggests to classify all difficult tumors as probably malignant when an expert opinion is not readily available [26]. Another option is to use newer IOTA models such as LR2 model which incorporates only 6 variables without CA-125 assessment [27] or the ADNEX model which uses 3 clinical and 6 ultrasound variables with CA-125 necessary to predict stage II–IV ovarian cancers [28]. For both methods basic IOTA training is essential as cutoff point for predicting malignancy has extremely strong influence on sensitivity and specificity of the tested diagnostic model.

In conclusion, this retrospective study has shown that all ultrasound-based prognostic methods performed well in the preoperative discrimination of adnexal masses. In the absence of an expert examiner's opinion the IOTA group models such as Simple Rules and LR1 could be preferably used by the less experienced sonographers.

Authors' contributions

JK, AC, MF, PP, NS and DB had full access to all the data in the study and take responsibility for the integrity of the data, the accuracy of the data analysis, and the decision to submit for publication. JK, AC, NS conceived, designed and obtained funding for the study. JK and AC performed all ultrasound examinations. JK, AC, DB, NS analyzed the data. JK, MF, PP drafted the manuscript. All authors interpreted the data,

critically revised the draft for important intellectual content, and gave final approval of the manuscript to be published.

REFERENCES

- American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Gynecology. Practice Bulletin No. 174: Evaluation and Management of Adnexal Masses. *Obstet Gynecol.* 2016; 128(5): e210–e226, doi: [10.1097/AOG.0000000000001768](https://doi.org/10.1097/AOG.0000000000001768), indexed in Pubmed: [27776072](https://pubmed.ncbi.nlm.nih.gov/27776072/).
- Hall JA. Adnexal masses: when to observe, when to intervene, and when to refer. *Obstet Gynecol.* 2010; 116(2 Pt 1): 440; author reply 440, doi: [10.1097/AOG.0b013e3181ea4f15](https://doi.org/10.1097/AOG.0b013e3181ea4f15), indexed in Pubmed: [20664410](https://pubmed.ncbi.nlm.nih.gov/20664410/).
- Liu JH, Zanotti KM. Management of the adnexal mass. *Obstet Gynecol.* 2011; 117(6): 1413–1428, doi: [10.1097/AOG.0b013e31821c62b6](https://doi.org/10.1097/AOG.0b013e31821c62b6), indexed in Pubmed: [21606754](https://pubmed.ncbi.nlm.nih.gov/21606754/).
- Carney ME, Lancaster JM, Ford C, et al. A population-based study of patterns of care for ovarian cancer: who is seen by a gynecologic oncologist and who is not? *Gynecol Oncol.* 2002; 84(1): 36–42, doi: [10.1006/gyno.2001.6460](https://doi.org/10.1006/gyno.2001.6460), indexed in Pubmed: [11748973](https://pubmed.ncbi.nlm.nih.gov/11748973/).
- Giede KC, Kieser K, Dodge J, et al. Who should operate on patients with ovarian cancer? An evidence-based review. *Gynecol Oncol.* 2005; 99(2): 447–461, doi: [10.1016/j.ygyno.2005.07.008](https://doi.org/10.1016/j.ygyno.2005.07.008), indexed in Pubmed: [16126262](https://pubmed.ncbi.nlm.nih.gov/16126262/).
- Mercado C, Zingmond D, Karlan B, et al. Quality of care in advanced ovarian cancer: The importance of provider specialty. *Gynecologic Oncology.* 2010; 117(1): 18–22, doi: [10.1016/j.ygyno.2009.12.033](https://doi.org/10.1016/j.ygyno.2009.12.033).
- Valentin L, Hagen B, Tingulstad S, et al. Comparison of 'pattern recognition' and logistic regression models for discrimination between benign and malignant pelvic masses: a prospective cross validation. *Ultrasound Obstet Gynecol.* 2001; 18(4): 357–365, doi: [10.1046/j.0960-7692.2001.00500.x](https://doi.org/10.1046/j.0960-7692.2001.00500.x), indexed in Pubmed: [11778996](https://pubmed.ncbi.nlm.nih.gov/11778996/).
- Moszynski R, Szperek D, Szubert S, et al. Analysis of false negative results of subjective ultrasonography assessment of adnexal masses. *Ginekol Pol.* 2013; 84(2): 102–107, doi: [10.17772/gp/1548](https://doi.org/10.17772/gp/1548), indexed in Pubmed: [23668055](https://pubmed.ncbi.nlm.nih.gov/23668055/).
- Stukan M, Dudziak M, Ratajczak K, et al. Usefulness of diagnostic indices comprising clinical, sonographic, and biomarker data for discriminating benign from malignant ovarian masses. *J Ultrasound Med.* 2015; 34(2): 207–217, doi: [10.7863/ultra.34.2.207](https://doi.org/10.7863/ultra.34.2.207), indexed in Pubmed: [25614393](https://pubmed.ncbi.nlm.nih.gov/25614393/).
- Kajser J, Gorp TV, Smet ME, et al. Are serum HE4 or ROMA scores useful to experienced examiners for improving characterization of adnexal masses after transvaginal ultrasonography? *Ultrasound in Obstetrics & Gynecology.* 2013; 43(1): 89–97, doi: [10.1002/uog.12551](https://doi.org/10.1002/uog.12551).
- Timmerman D, Valentin L, Bourne TH, et al. International Ovarian Tumor Analysis (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.* 2000; 16(5): 500–505, doi: [10.1046/j.1469-0705.2000.00287.x](https://doi.org/10.1046/j.1469-0705.2000.00287.x), indexed in Pubmed: [11169340](https://pubmed.ncbi.nlm.nih.gov/11169340/).
- Timmerman D, Testa AC, Bourne T, et al. International Ovarian Tumor Analysis Group. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group. *J Clin Oncol.* 2005; 23(34): 8794–8801, doi: [10.1200/JCO.2005.01.7632](https://doi.org/10.1200/JCO.2005.01.7632), indexed in Pubmed: [16314639](https://pubmed.ncbi.nlm.nih.gov/16314639/).
- Timmerman D, Testa AC, Bourne T, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. *Ultrasound Obstet Gynecol.* 2008; 31(6): 681–690, doi: [10.1002/uog.5365](https://doi.org/10.1002/uog.5365), indexed in Pubmed: [18504770](https://pubmed.ncbi.nlm.nih.gov/18504770/).
- Polish Gynecological Society recommendations. *Ginekol.Pol.Supl.* 2015.
- 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(10): 922–929, doi: [10.1111/j.1471-0528.1990.tb02448.x](https://doi.org/10.1111/j.1471-0528.1990.tb02448.x), indexed in Pubmed: [2223684](https://pubmed.ncbi.nlm.nih.gov/2223684/).
- Tingulstad S, Hagen B, Skjeldestad FE, et al. 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.* 1996; 103(8): 826–831, doi: [10.1111/j.1471-0528.1996.tb09882.x](https://doi.org/10.1111/j.1471-0528.1996.tb09882.x), indexed in Pubmed: [8760716](https://pubmed.ncbi.nlm.nih.gov/8760716/).
- Tingulstad S, Hagen B, Skjeldestad FE, et al. The risk-of-malignancy index to evaluate potential ovarian cancers in local hospitals. *Obstet Gynecol.* 1999; 93(3): 448–452, doi: [10.1097/00006250-199903000-00028](https://doi.org/10.1097/00006250-199903000-00028), indexed in Pubmed: [10074998](https://pubmed.ncbi.nlm.nih.gov/10074998/).
- Amor F, Alcázar JL, Vaccaro H, et al. GI-RADS reporting system for ultrasound evaluation of adnexal masses in clinical practice: a prospective multicenter study. *Ultrasound Obstet Gynecol.* 2011; 38(4): 450–455, doi: [10.1002/uog.9012](https://doi.org/10.1002/uog.9012), indexed in Pubmed: [21465605](https://pubmed.ncbi.nlm.nih.gov/21465605/).
- Amor F, Alcázar JL, Vaccaro H, et al. GI-RADS reporting system for ultrasound evaluation of adnexal masses in clinical practice: a prospective multicenter study. *Ultrasound Obstet Gynecol.* 2011; 38(4): 450–455, doi: [10.1002/uog.9012](https://doi.org/10.1002/uog.9012), indexed in Pubmed: [21465605](https://pubmed.ncbi.nlm.nih.gov/21465605/).
- Education and Practical Standards Committee, European Federation of Societies for Ultrasound in Medicine and Biology. Minimum training recommendations for the practice of medical ultrasound. *Ultraschall Med.* 2006; 27(1): 79–105, doi: [10.1055/s-2006-933605](https://doi.org/10.1055/s-2006-933605), indexed in Pubmed: [16508866](https://pubmed.ncbi.nlm.nih.gov/16508866/).
- Miller RW, Ueland FR. Risk of malignancy in sonographically confirmed ovarian tumors. *Clin Obstet Gynecol.* 2012; 55(1): 52–64, doi: [10.1097/GRF.0b013e31824970cf](https://doi.org/10.1097/GRF.0b013e31824970cf), indexed in Pubmed: [22343229](https://pubmed.ncbi.nlm.nih.gov/22343229/).
- 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(12): 2448–2454, doi: [10.1038/bjc.2013.224](https://doi.org/10.1038/bjc.2013.224), indexed in Pubmed: [23674083](https://pubmed.ncbi.nlm.nih.gov/23674083/).
- Knafel A, Banas T, Nocun A, et al. The Prospective External Validation of International Ovarian Tumor Analysis (IOTA) Simple Rules in the Hands of Level I and II Examiners. *Ultraschall Med.* 2016; 37(5): 516–523, doi: [10.1055/s-0034-1398773](https://doi.org/10.1055/s-0034-1398773), indexed in Pubmed: [26126150](https://pubmed.ncbi.nlm.nih.gov/26126150/).
- Abdalla N, Winiarek J, Bachanek M, et al. Clinical, ultrasound parameters and tumor marker-based mathematical models and scoring systems in pre-surgical diagnosis of adnexal tumors. *Ginekol Pol.* 2016; 87(12): 824–829, doi: [10.5603/GP.2016.0096](https://doi.org/10.5603/GP.2016.0096), indexed in Pubmed: [28098934](https://pubmed.ncbi.nlm.nih.gov/28098934/).
- Valentin L, Amey L, Savelli L, et al. Adnexal masses difficult to classify as benign or malignant using subjective assessment of gray-scale and Doppler ultrasound findings: logistic regression models do not help. *Ultrasound in Obstetrics & Gynecology.* 2011; 38(4): 456–465, doi: [10.1002/uog.9030](https://doi.org/10.1002/uog.9030).
- Kajser 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(3): 449–462, doi: [10.1093/humupd/dmt059](https://doi.org/10.1093/humupd/dmt059), indexed in Pubmed: [24327552](https://pubmed.ncbi.nlm.nih.gov/24327552/).
- Meys EMJ, Kajser J, Kruitwagen RF, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis. *Eur J Cancer.* 2016; 58: 17–29, doi: [10.1016/j.ejca.2016.01.007](https://doi.org/10.1016/j.ejca.2016.01.007), indexed in Pubmed: [26922169](https://pubmed.ncbi.nlm.nih.gov/26922169/).
- Nunes N, Ambler G, Foo X, et al. A prospective evaluation of the IOTA Logistic Regression Models (LR1 and LR2) in comparison to Subjective Pattern Recognition for the diagnosis of ovarian cancer in the outpatient setting. *Ultrasound Obstet Gynecol.* 2017 [Epub ahead of print], doi: [10.1002/uog.18918](https://doi.org/10.1002/uog.18918), indexed in Pubmed: [28976616](https://pubmed.ncbi.nlm.nih.gov/28976616/).
- Calster BV, Hoorde KV, Valentin L, et al. Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study. *BMJ.* 2014; 349(oct07 3): g5920–g5920, doi: [10.1136/bmj.g5920](https://doi.org/10.1136/bmj.g5920).
- Abramowicz JS, Timmerman D. Ovarian mass-differentiating benign from malignant: the value of the International Ovarian Tumor Analysis ultrasound rules. *Am J Obstet Gynecol.* 2017; 217(6): 652–660, doi: [10.1016/j.ajog.2017.07.019](https://doi.org/10.1016/j.ajog.2017.07.019), indexed in Pubmed: [28735703](https://pubmed.ncbi.nlm.nih.gov/28735703/).