# Performance of ultrasound models in diagnosis of Ovarian Cancer: Experience of a Portuguese tertiary center

# Performance de modelos ecográficos na predição do Cancro do Ovário: Experiência de um centro terciário Português

Ângela Rodrigues[1], Liana Negrão[2], Fernanda Águas[3], Maria Geraldina Castro[4]
Centro Hospitalar e Universitário de Coimbra, Departamento de Ginecologia

## Abstract

**Overview and Aims:** The purpose of this study was to find the best ultrasound model for preoperative discrimination between benign and malignant adnexal masses in a group of Portuguese women.
**Methods:** Single-centre retrospective study of 123 adnexal masses. The ultrasound images were described by an experienced ultrasonographer and classified as benign or malignant, according to IOTA simple rules (SR), RMI scoring and logistic regression model L2 (LR2). Two study groups were considered according to histologic diagnosis (benign and malignant). Borderline tumours were counted as malignant. The sensitivity, specificity, positive and negative predictive values (PPV and NPV) were calculated for IOTA SR; LR2 model; RMI score and use of serum CA 125 as a secondstage test in cases of IOTA SR complemented by subjective assessment.
**Results:** Among the 123 tumours, 81.3% were benign and 18.7% were malignant on histolopathology. When inconclusive tumors were considered malignant, the IOTA SR had a sensitivity of 95.6%, specificity of 69.9%, PPV 46.8% and NPV of 98.3%. If inconclusive tumors were classified by subjective sonographic assessment, IOTA SR had a sensitivity of 91.3% and specificity of 78.3%. The LR2 model had a sensitivity of 91.3%, specificity 77.1%, PPV 63.6% and NPV 93.06%.
**Conclusion:** IOTA SR and IOTA LR2 models achieved the best diagnostic accuracy for differentiating between benign or malignant adnexal masses. In case of inconclusive results, subjective assessment of ultrasound findings by expert examiners should be incorporated.

**Keywords:** Adnexal masses; Ovarian cancer; IOTA simple rules; CA 125 level; Ultrasound diagnostic accuracy

## INTRODUCTION

Globally, ovarian cancer is the seventh most prevalent cancer in women and the eighth most deadly[1]. In Europe, there were 45 994 new cases during 2018, with an age-adjusted incidence rate of 15.7 per 100 000 women[2]. In Portugal, in 2018, the incidence was of 7.0 per 100 000 women[2] and the standardised mortality rate for all ages was 4.3 per 100 000 women[2].

The five-year survival rates are below 45%, but can reach 90% when diagnosis is made at stage I of the disease[1]. However, only 15% of ovarian cancers are detected in this stage.

The benign or malignant preoperative diagnosis of an adnexal mass is critical to plan patient management and optimize treatment. Ultrasound is the best modality for differentiating between a benign or malignant adnexal mass[3]. Many scoring systems for malignancy risk assessment of an ovarian tumour based on ultrasound findings have been proposed to improve diagnostic performance. However, the best ultrasound model to apply in clinical practice is not yet fully defined.

The International Ovarian Tumor Analysis (IOTA) group has incorporated ultrasound features into its Logistic Regression model 2 (LR2) and its Simple Rules model that characterize adnexal tumours based upon the presence or absence of typical ultrasound features

1. Interna de formação específica de Ginecologia e Obstetrícia, Centro Hospitalar e Universitário de Coimbra- CHUC; Assistente Convidada da Faculdade de Medicina da Universidade Coimbra
2. Assistente Graduado Hospitalar de Ginecologia e Obstetrícia do Centro Hospitalar e Universitário de Coimbra
3. Assistente Graduada Sénior, Diretora do Serviço de Ginecologia, Centro Hospitalar e Universitário de Coimbra - CHUC
4. Assistente Hospitalar de Ginecologia e Obstetrícia do Centro Hospitalar e Universitário de Coimbra

of malignancy[4,5,7]. The risk of malignancy index (RMI)[5] was the most widely used model, but the IOTA group studies demonstrated that IOTA models have higher diagnostic performance in ovarian malignancy[4,7].

Few studies originating from groups outside the IOTA group have been developed to compare the pre-operative diagnostic accuracy of the various ultrasound scores, including in Portugal.

Therefore, finding the best ultrasound model to predict ovarian malignancy in a Portuguese population and consequently referring these patients to gynaecological oncologists is the main objective of this study.

We compared the diagnostic accuracy of ultrasound--based simple rules of *Timmerman at al.*[4] alone or in association with serum CA 125 level, LR2 and RMI for the pre-operative discrimination between benign and malignant adnexal masses in Portuguese women.

## MATERIALS AND METHODS

This was a single-centre retrospective study and a diagnostic accuracy study conducted in a Portuguese tertiary referral center for gynaecologic oncology. We included 123 patients who underwent an elective surgical treatment for an adnexal tumour at our hospital. Patient selection was made in a comprehensive database research and after informed consent signature to participate in the study. The database is blinded relative to the patient's identification in order to prevent investigators from recognising the subjects and in order to maintain confidentiality. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki and approved by the local ethics committee.

We collected information regarding age, body mass index (BMI) (kg/m$^2$), menarche age, menopausal status, parity, current or previous hormonal contraception or menopausal therapy, family or personal history of breast cancer or ovarian cancer and serum CA 125 level.

### Ultrasound examination

In most cases included, prior to surgery, ultrasound evaluations were performed in our centre by a single gynaecologist with 10 years' experience in gynaecological ultrasound and with European IOTA certification (M.G.C.). A minority of ultrasound images were performed by other three experienced sonographers of the department and described by M.C.G. without knowledge of the histological outcome, patient's CA 125 serum levels and preoperative ultrasound report.

All women performed transvaginal grayscale and colour Doppler ultrasound examination using a Voluson E8 ultrasound machine (GE Healthcare Ultrasound). An additional transabdominal scan was conducted if the mass was too large or if malignancy was suspected.

In all cases, sonographic features of adnexal tumours were described using terms and definitions as published by the IOTA group[5]. The ultrasound examiner classified each mass as benign or malignant according IOTA simple rules, RMI scoring and logistic regression model L2 (LR2).

We applied the simple rules as reported by Timmerman at al[4]. We used five ultrasound features to predict a malignant tumour (M-features): irregular solid tumour (M1), ascites (M2), at least four papillary structures (M3), irregular multilocular solid tumour with a largest diameter of 100 mm or more (M4) and high Doppler flow (colour score 4) (M5). We also used five ultrasound benign features to predict a benign tumour (B-features): unilocular ovarian cyst (B1), presence of solid components smaller than 7 mm in diameter (B2), presence of acoustic shadows (B3), smooth multilocular tumour with largest diameter < 10 cm (B4) and no detectable colour Doppler signal (B5). If one or more M-features were present in the absence of a B-feature, the mass was categorized as malignant. If one or more B-features applied in the absence of an M-feature, we classified the mass as benign. If both or no M-features and B-features were present, the mass was categorized as inconclusive. In this case, we adopted two strategies. First, the tumour was classified subjectively as benign or malignant by the experienced ultrasonographer and, second, all inconclusive tumours were considered malignant.

The LR2 model uses six predictors[7]. The variables included were: (1) age of the patient (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) maximal diameter of the solid component (expressed in mm but with no increase above 50 mm); (5) irregular internal cyst walls (yes = 1, no = 0); and (6) the presence of acoustic shadows (yes = 1, no = 0). An adnexal tumour with an estimated probability of malignancy > 0.10 by LR2 was classified as malignant.

The RMI was determined using the ultrasound features, serum CA 125 level and menopausal status. The ultrasound score (U) was calculated for each patient considering five ultrasound features suggestive of can-

cer: multilocular cyst, solid areas, bilateral masses, ascites and intra-abdominal metastases. U was assigned a value of 0, 1 or 3 if none, one or two or more of these features were present, respectively. A score (M) of 1 was assigned to premenopausal women and 3 to postmenopausal women. As described by *Jacobs et al.*[6], RMI scoring was calculated as follows: RMI score = U x M x serum CA 125 (U/mL). An adnexal mass with a score ≥ 200 was classified as malignant.

The ultrasound models were calculated using a secure electronic data-collection system (Astraia version 1.25.2, Astraia Software).

### Surgery and pathology analysis

Treatment surgery was performed at our centre and the surgical procedures and techniques were chosen according to surgeon experience and clinical indication.

Histopathology diagnosis of surgical specimens was the clinical reference used for all patients in this study. The pathologic study was performed in the department of Pathologic Anatomy of our hospital, following the guidelines of the World Health Organization International Classification of Ovarian Tumours[8]. All masses were classified into two groups: benign or malignant.

### Statistical analysis

Statistical analysis was performed using SPSS version 21.0. All statistical calculations were performed using 95% CIs and considering $p < 0.05$ as statistically significant.

We considered two study groups according to histologic diagnosis (benign and malignant). For statistical purposes, borderline tumours were counted as malignant.

Categorical variables were described as frequencies and percentages and continuous variables as means and standard deviations.

We compared the clinical and demographic parameters of groups using the chi-square for categorical variables and independent sample T-test for continuous variables. A binary logistic regression model was used to determine the ultrasound characteristics of adnexal masses with greater impact on the prediction of malignancy. The logistic regression was adjusted for age and menopausal status.

We generated a ROC curve to determine the point at which serum level CA-125 had the highest sensitivity and specificity in distinguishing preoperatively between benign and malignant adnexal masses. The area under the ROC curve (AUC) was calculated, and 95% confidence intervals were used.

The sensitivity, specificity, positive and negative predictive values (PPV and NPV) were calculated for the following six diagnostic models: (1) IOTA Simple Rules (if inconclusive, considered malignant), (2) IOTA simple Rules (if inconclusive, complemented by subjective evaluation), (3) LR2 model, (4) RMI score and use of serum CA 125 as a second-stage test in cases of IOTA simples rules complemented by subjective evaluation predicts a benign tumour (5) or malignant tumour (6). The CA-125 cut-off used in the last two models (5 and 6) was calculated by the ROC curve. In model 5, benign masses with CA 25 above the cut-off were considered malignant and in model 6 malignant masses with CA 125 below the cut-off were considered benign to evaluate the diagnostic performance of these strategic diagnostics (Figure 1).

## RESULTS

The mean age of the sample was 46.98 ± 15.82 years (range 17-86 years). Most women (65.90%) were premenopausal and 23.90% were nulliparous. In 18.90% and 1.60% of patients there was a family history of breast and ovarian cancer, respectively.

Of the 123 women recruited in the one year time period with suspicious adnexal masses, four had a bilateral tumour. In the case of bilateral adnexal masses, we included the mass with the most complex ultrasonic morphology in our statistical analysis. If both masses had similar ultrasonic morphology, we included the largest one or the one most easily accessible by transvaginal ultrasound. Table I shows the histological diagnosis of the masses. Of these, 100 (81.3%) were benign and 23 (18.7%) were malignant. Among all benign lesions, serous cystadenoma (35%) and mucinous cystadenoma (21%) were the most frequent histological types. In the malignant group, serous cystadenocarcinoma (60.9%) was the most prevalent and 26% of the malignancies were borderline tumours.

The mean age was significantly higher in the malignant group, (54.48 ± 14.76 vs 45.26 ± 15.62 years, $p = 0.01$) and postmenopausal women were significantly more prevalent in this group (30.0% vs 52.2%, $p = 0.04$). Use of hormonal contraception or menopausal therapy, age at menarche, parity and family history of cancer were not significantly different be-

**TABLE I. HISTOLOGY OF THE ADNEXAL MASSES ACCORDING TO WHO HISTOLOGIC CLASSIFICATION (2014)**

| Benign histology | n(%) | Malignant histology | n(%) |
|---|---|---|---|
| Serous cystadenoma | 35 (35.0) | Serous cystadenocarcinoma | 14 (60.9) |
| Mucinous cystadenoma | 21 (21.0) | Seromucinous carcinoma | 1 (4.3) |
| Teratoma | 17 (17.0) | Undifferentiated carcinoma | 1 (4.3) |
| Serous surface papilloma | 9 (9.0) | Carcinosarcoma | 1 (4.3) |
| Fibroma | 7 (7.0) | Seromucinous borderline tumour | 3 (13.0) |
| Serous adenofibroma | 3 (3.0) | Mucinous borderline tumour | 2 (8.7) |
| Fibrothecoma | 3 (3.0) | Serous borderline tumour | 1 (4.3) |
| Seromucinous cystadenoma | 3 (3.0) | | |
| Brenner tumour | 2 (2.0) | | |
| Benign (total) | 100 (81.3) | Malignant (total) | 23 (18.7) |

**TABLE II. CLINICAL CHARACTERISTICS OF WOMEN WITH ADNEXAL TUMOURS ACCORDING TO HISTOLOGICAL DIAGNOSIS GROUPS**

| | Histological diagnosis | | |
|---|---|---|---|
| Clinical characteristics | Benign (n = 100) | Malignant (n = 23) | p |
| Age (years, mean ± SD) | 45.26 ± 15.62 | 54.48 ± 14.76 | 0.01 |
| ≤ 20 years (n%) | 6 (6.0%) | 0 | 0.01 |
| 21-50 years (n%) | 60 (60.0%) | 9 (39.1%) | |
| > 50 years (n%) | 34 (34.0%) | 14 (60.9%) | |
| Body mass index (kg/m2, mean ± SD) | 25.03 ± 3.73 | 27.40 ± 7.10 | 0.04 |
| Menarche age (years, mean ± SD) | 12.36 ± 1.33 | 12.15 ± 0.75 | 0.51 |
| Menopausal status | | | 0.04 |
| Premenopausal (n%) | 70 (70.0%) | 11 (47.8%) | |
| Postmenopausal (n%) | 30 (30.0%) | 12 (52.2%) | |
| Parity | | | 0.779 |
| Nulliparous (n%) | 24 (24.7%) | 4 (20.0%) | |
| Multiparous (n%) | 73 (75.3%) | 16 (80.0%) | |
| Current or previous hormonal contraception (n%) | 79 (79.8%) | 18 (81.8%) | 0.547 |
| Current or previous menopausal hormonal therapy (n%) | 5 (5.1%) | 3 (13.6%) | 0.158 |
| Family history of breast cancer and/or ovarian cancer (n%) | 17 (17.2%) | 7 (30.4%) | 0.156 |
| Degree of relative with family history of breast/ovarian cancer | | | 0.05 |
| First degree | 9 (52.9%) | 7 (100%) | |
| Second degree | 8 (47.05%) | 0 | |

tween groups (Table II).

Concerning CA-125, the cut-off value with highest sensitivity and specificity in differentiating malignant from benign ovarian tumors determined by ROC curve was 23.5 U/mL (area under the curve, 0.872 (95% IC, 0.782 – 0.961) – Figure 2. The cut-off sensitivity was 85.0%, specificity was 70.7%, PPV was 43.6% and NPV was 94.6%. In the malignant group, 85.0% of women had CA 125 levels ≥ 23.5 U/ml. In this group, three had CA-125 levels < 23.5 U/mL, two were mucinous borderline tumours and one a low-degree serous cystoadenocarcinoma (stage I).

Regarding ultrasound findings, malignant tumours had significantly larger diameter ($p < 0.01$), greater vascularization ($p < 0.01$) and were predominantly solid ($p < 0.01$). Table III shows ultrasound characteristics of the adnexal masses (excluding teratomas) between the study groups. In the unadjusted binary logistic regression model (Table IV), the score of 4 on vascularization ($p < 0.01$), the largest tumour at least 10 cm ($p < 0.01$) and the solid composition were statistically associated with higher risk of malignancy. However, in

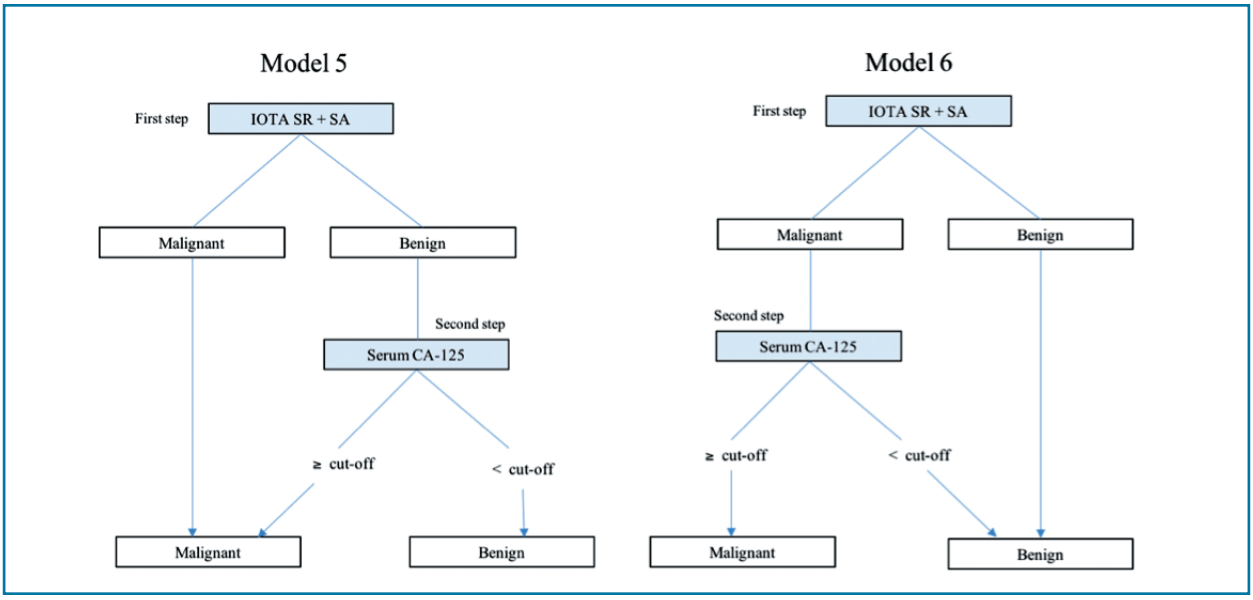


**FIGURE 1.** Models illustrating the use of serum CA 125 as a second-stage test in cases where IOTA simple rules complemented by subjective assessment in inconclusive masses predicts a malignant (model 5) and a benign (model 6) tumour (adapted from the study by Lil Vicent el al.[17]). Abbreviations: SR – simple rules; SA – subjective assessment

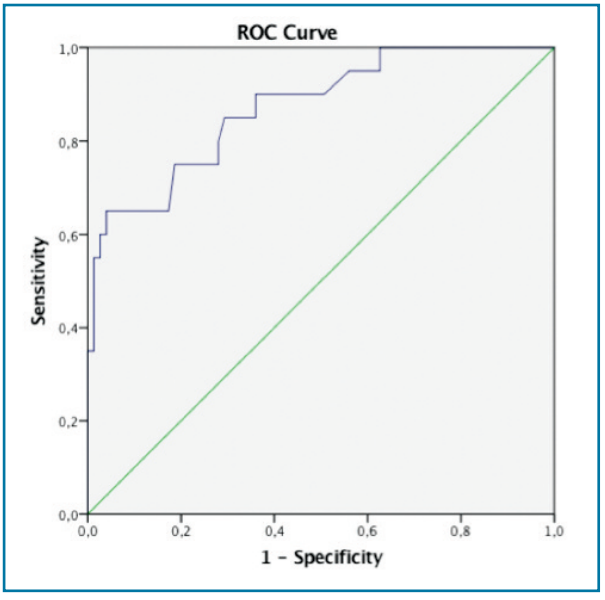


**FIGURE 2.** ROC curve of CA-125 performance in differentiating malignant from benign ovarian tumours.
AUC= 0.872 (95% IC, 0.782 – 0.961)

analysis adjusted to patient age and menopausal status, these three ultrasound characteristics were independently associated with risk of malignancy of the adnexal tumors, with the score of 4 on vascularization being the one with the highest association.

IOTA simple rules could be applied to 106 (86.2%) tumors. The other 17 tumors correspond to teratomas and their diagnosis was made by pattern recognition. Among these 106 tumors, this ultrasound model was inconclusive in 24 (22.6%). For the inconclusive results by IOTA simple rules, two strategies were applied. In the first (all inconclusive tumors considered malignant), the IOTA simple rules had a sensitivity of 95.6% (95% IC, 76.0-99.8), specificity of 69.9% (95% IC, 58.7-79.2), PPV 46.8% and NPV of 98.3%. In the second, the inconclusive tumors were classified subjectively as benign or malignant by an experienced ultrasonographer and the IOTA simple rules had a sensitivity of 91.3% (95% IC, 70.5-98.5), specificity of 78.3% (95% IC, 67.6-86.3), PPV 53.8% and NPV 97% (Table V). The LR2 model had a sensitivity of 91.3% (95% IC, 70.5-98.5), specificity 77.1% (95% IC, 66.3--85.3), PPV 63.6% and NPV 93.06%. The sensitivity of RMI score (73.7%) was significantly lower than IOTA models.

Of the 123 adnexal masses, 85 had available CA 125 results. Therefore, 85 adnexal masses were included in performance evaluation of the strategic diagnostic, when serum CA 125 was used as a second-stage test of the IOTA simple rules complemented by subjective evaluation. Whenever CA-125>23.5U/mL predicts the

**TABLE III. ULTRASOUND CHARACTERISTICS OF THE 106 ADNEXAL TUMOURS (EXCLUDING TERATOMAS) BETWEEN GROUPS**

| Clinical characteristics | Histological diagnosis | | p |
|---|---|---|---|
| | Benign (n = 83) | Malignant (n = 23) | |
| Larger diameter (cm, mean ±SD) | 7.2±4.6 | 10.4±5.4 | 0.005 |
| <10 cm | 81.9% | 47.8% | 0.001 |
| ≥ 10 cm | 18.1% | 52.2% | |
| Laterality | | | |
| Unilateral | 95.2% | 100% | 0.575 |
| Bilateral | 4.8% | 0% | |
| Locularity | | | |
| Unilocular | 50.0% | 29.4% | 0.121 |
| Multilocular | 50% | 70.6% | |
| Composition | | | |
| Solid | 18.1% | 69.6% | |
| Cystic | 53.0% | 0% | 0.000 |
| Solid + cystic | 28.9% | 30.4% | |
| Degree of vascularization | | | |
| Absent | 56.6% | 4.3% | |
| Minimum | 24.1% | 17.4% | 0.000 |
| Moderate | 15.7% | 47.8% | |
| Intense | 3.6% | 30.4% | |

**TABLE IV. EFFECT OF ULTRASOUND CHARACTERISTICS OF THE ADNEXAL MASSES IN THE PREDICTION OF MALIGNANCY (UNADJUSTED AND ADJUSTED FOR AGE AND MENOPAUSAL STATUS)**

| | Logistic Regression | | | |
|---|---|---|---|---|
| | Unadjusted | | Adjusted* | |
| Ultrasound characteristics | OR (95% CI) | p | OR (95% CI) | p |
| Largest tumour diameter ≥ 10 cm | 4.94 (1.84-13.32) | 0.000 | 5.29 (1.90-14.81) | 0.001 |
| Solid composition | 3.66 (1.22-10.96) | 0.021 | 3.62 (1.17-11.20) | 0.026 |
| Vascularization score 4 | 11.67 (2.72-49.99) | 0.000 | 10.19 (2.23-46.63) | 0.003 |

*Logistic regression adjusted for age and menopausal status

malignancy of a mass classified by IOTA simple rules complemented by subjective assessment as benign (Figure 1 – Model 5), the sensibility increases to 100%, with a significant specificity reduction to 53.8% and the false positive rate increases by 24.7% (Table VI), compared to IOTA simple rules supplemented by subjective evaluation alone (Table V). When CA-125 < 23.5U/mL predicts the benignity of a mass classified by IOTA simple rules as malign (Figure 1 – Model 6), the sensibility decreases and the false negative rate increases by 21.3% (Table VI), compared to IOTA simple rules supplemented by subjective evaluation alone (Table V).

## DISCUSSION

In this study, the performance of IOTA simple rules was excellent and was more effective in the differentiation between benign and malignant tumours. This finding was similar to that found in the original study by Timmerman et al.[9]. The prevalence of classifiable tumours with IOTA simple rules in the original research was 77% vs 86.2% in our study. When inconclusive adnexal masses by IOTA simple rules were classified according to subjective assessment of ultrasound findings by an ultrasound examiner in a second-stage test, performance was as follows: sensitivity 91% in the original research vs 91.3% in our study and specificity 93% vs 78.3%. When all inconclusive tumours were

**TABLE V. PERFORMANCE OF DIAGNOSTIC TESTS**

| | Histologic diagnosis | | | Performance | | | |
|---|---|---|---|---|---|---|---|
| Diagnostic tests | Benign n(%) | Malignant n(%) | p | Sensitivity (% (95% CI)) | Specificity (% (95% CI)) | PPV (%) | NPV (%) |
| IOTA simple rules (if inconclusive, supplemented with SA) | | | | | | | |
| Benign | 78.3% | 8.7% | 0.000 | 91.3% (70.5-98.5) | 78.3% (67.6-86.3) | 53.8% | 97.0% |
| Malignant | 21.7% | 91.3% | | | | | |
| IOTA simple rules (if inconclusive, considered malignant) | | | | | | | |
| Benign | 69.9% | 4.3% | 0.000 | 95.6% (76.0-99.8) | 69.9% (58.7-79.2) | 46.8% | 98.3% |
| Malignant | 30.1% | 95.7% | | | | | |
| IOTA Logistic regression model 2 (LR2) | | | | | | | |
| < 0.10 | 77.1% | 8.7% | 0.000 | 91.3% (70.5-98.5) | 77.1% (66.3-85.3) | 52.5% | 97.0% |
| > 0.10 | 22.9% | 91.3% | | | | | |
| RMI score | | | | | | | |
| Benign | 89.3% | 26.3% | 0.000 | 73.7% (48.6-89.9) | 89.3 % (79.5-95.0) | 63.6% | 93.06% |
| Malignant | 10.7% | 73.7% | | | | | |

Abbreviations: NPV, negative predictive value; PPV, positive predictive value; RMI, risk of malignancy index; SA, subjective assessment

**TABLE VI. DIAGNOSTIC PERFORMANCE WHEN USING SERUM CA 125 (CUT-OFF 23.5 U/ML) AS A SECOND-STAGE TEST AFTER IOTA SIMPLE RULES COMPLEMENTED BY SUBJECTIVE EVALUATION IN THE INCONCLUSIVE MASSES**

| | Histologic diagnosis | | | Performance | | | |
|---|---|---|---|---|---|---|---|
| | Benign n(%) | Malignant n(%) | p | Sensitivity (% (95% CI)) | Specificity (% (95% CI)) | PPV (%) | NPV (%) |
| IOTA simple rules (supplemented with SA) + serum CA 125§ | | | | | | | |
| Benign | 53.8% | 0% | 0.000 | 100% (79.1-100) | 53.8% (41.11-66.1) | 38.1% | 100% |
| Malignant | 46.2% | 100% | | | | | |
| IOTA simple rules (supplemented with SA) + serum CA 125§§ | | | | | | | |
| Benign | 87.7% | 30.0% | 0.000 | 70.0% (45.7-87.2) | 87.7% (76.6-94.2) | 63.6% | 90.5% |
| Malignant | 12.3% | 70% | | | | | |

Abbreviations: NPV, negative predictive value; PPV, positive predictive value; SA, subjective assessment
§ benign masses classified by IOTA simples rules (supplemented by SA) with CA 25 above the cut-off 23.5 U/mL were considered malignant (model 5)
§§ malignant masses classified by IOTA simples rules (supplemented by SA) with CA 25 below the cut-off 23.5 U/mL were considered benign (model 6)

considered malignant, the sensitivity was 95% *vs* 95.6% and the specificity 78% *vs* 69.9%, in the original and our study, respectively. Comparison of our findings with studies of other groups is limited. A few studies have been conducted to test the performance of the simple rules and to compare it with other ultrasound tests, particularly the LR2 model and RMI score. Alcázar et al. (2013) conducted a prospective external validation of the IOTA simple rules in a series of 340 masses. The simple rules were applicable in 79.4% of the tumours. Sensitivity and specificity were 87.9% and 97.5%, respectively. In inconclusive results, the sensi-

tivity and specificity obtained by the expert examiner were 90.9% and 89.6%, respectively[10]. Tantipalakorn et al. (2014) performed a prospective study using the IOTA simple rules on a series of 398 adnexal masses. This model was applicable in 80.1% and the sensitivity of the simples rules was 82.9% and specificity 95.3%[11]. More recently, Koneczny et al. (2017) compared different ultrasound models in 271 masses. Simple rules were applicable in 87% and had sensitivity of 90.6% with specificity of 95.3%. RMI had a sensitivity of 55.3% with a specificity of 94%[12]. In a meta-analysis study, Meys et al. (2016) analysed 47 prospective diagnostic studies, including 19 674 adnexal tumours. The simple rules (classifying inconclusives as malignant) had a pooled sensitivity and specificity, respectively, of 0.93 [95% CI 0.91-0.95] and 0.80 [95% CI 0.77-0.82]. Simple rules in combination with subjective assessment for inconclusive results gave only small differences (sensitivity 0.91 [95% CI 0.89-0.93] and specificity 0.91 [95% CI 0.87-0.94]). LR2 had comparable sensitivity and specificity to the strategy using simple rules. The RMI showed the worst diagnostic performance in comparison to other methods (sensitivity 0.75 [95% CI 0.72-0.79], specificity 0.92 [95% CI 0.88-0.94])[13]. These results were similar to a previous systematic review and meta-analysis[14] and a recent prospective study[15]. The last two studies[13,14] also demonstrated that both IOTA simple rules and IOTA LR2 model were the best diagnostic tests, with similar performance.

According to our results, the simple rules provide a accurate test to discriminate between benign and malignant adnexal lesions. They were applicable in 86.2% of tumours and the diagnostic performance in terms of sensitivity was similar to the original IOTA study[9] and others prospective studies included in meta-analysis[13,14], but the specificity was slightly lower. When simple rules are inconclusive, our results and others studies[13,14], show that a subjective assessment of ultrasound findings by expert examiners should be integrated[18]. The score of 4 on vascularisation of tumour ($p = 0.003$), the largest tumour diameter greater than at least 10 cm ($p = 0.001$) were the ultrasound features with the highest prognostic power for malignancy.

In our study, the IOTA LR2 model had the same diagnostic accuracy as IOTA simple rules supplemented by expert examination in inconclusive tumours (Sensibility LR2 model 91.3% vs IOTA simple rules 91.3%). The same result was shown in other studies[13,14].

The RMI classification in this study had a particularly low sensitivity in the diagnosis of adnexal masses (73.7% [95% CI 0.88-0.94]), as well as in other systematic review and meta-analysis[13,14]. The disadvantage of RMI is that it depends mainly on the level of serum CA-125, which has a limited value in the diagnosis of ovarian cancer, especially in premenopausal women. Various clinical conditions are associated with an increase in this marker, such as pregnancy, menses, endometriosis, pelvic inflammatory disease, fibroids and surgery[14,16-18]. Furthermore, in early stages of ovarian cancer, CA-125 is not increased. Consequently, in our study this biomarker (cut-off level 23.5 U/mL) did not improve the diagnostic performance of IOTA simple rules complemented by subjective assessment. However, it is important to point out that these results represent ultrasound evaluations performed by an experienced ultrasonographer. We do not know if the same results can be extrapolated to the less experienced ultrasonographers. This result was similar to Valentin L. et al.'s study, which considered that an ultrasound examination interpreted by an experienced operator in inconclusive masses is superior to the complementary analysis with serum CA-125[19].

The strength of this study is the comparison between different ultrasound models on a sample of Portuguese patients. The mains limitations were the reduced sample size and the retrospective character of the study.

In conclusion, our results showed that the IOTA simple rules and IOTA LR2 model achieve the best diagnostic accuracy for differentiating between a benign or malignant adnexal mass. Simple rules are an easy model to use in clinical practice. However, in cases of inconclusive results (in approximately 20%), subjective assessment of ultrasound findings by expert examiners should be incorporated . If such expertise is not available, the LR2 model can be an option.

## REFERENCES

1. Webb PM, Jordan SJ. Epidemiology of epithelial ovarian cancer. Best Pract Res Clin Obstet Gynaecol 2017. Doi: 10.1016/j.bpobgyn.2016.08.006.

2. Ferlay J, Colombet M, Soerjomataram I, Dyba T, Randi G, Bettio M, Gavin A, Visser O, Bray F. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. Eur J Cancer 2018. DOI: 10.1016/j.ejca.2018.07.005

3. Fischerova D, Burgetova A. Imaging techniques for the evaluation of ovarian cancer. Best Pract Res Clin Obstet Gynaecol 2014. Doi: 10.1016/j.bpobgyn.2014.04.006

4. Timmerman D, Testa AC, Bourne T, Ameye L, Jurkovic D,

Van Holsbeke C, Paladini D, Van Calster B, Vergote I, Van Huffel S, Valentin L. Simple ultrasound-based rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol 2008. DOI: 10.1002/uog.5365

5. Timmerman D, Velentin L, Bourne TH, Collins WP, Verrelst H, Vergote I. Terms, definitions and measurements to describe to sonographic features of adnexal tumors: a consensus opinion from the international ovarian tumor analysis (IOTA) group. Ultrasound Obstet Gynecol 2000; DOI: 10.1046/j.1469-0705.2000.00287.x.

6. Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas JG. 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. DOI: 10.1111/j.1471-0528.1990.tb02448.x

7. Timmerman D, Testa AC, Bourne T, Ferrazzi E, Ameye L, Konstantinovic ML, Van Calster B, Collins WP, Vergote I, Van Huffel S, Valentin L; International Ovarian Tumor Analysis Group. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicentre study by the International Ovarian Tumor Analysis Group. J Clin Oncol 2005; DOI: 10.1200/JCO.2005.01.7632

8. Kurman RJ, Carcangiu ML, Herrington CS. WHO Classification of Tumors of the Female Reproductive Organs. Lyon: WHO Press, 2014.

9. Timmerman, D., Ameye, L., Fischerova, D., Epstein, E., Melis, G. B., Guerriero, S., Van Holsbeke, C., Savelli, L., Fruscio, R., Lissoni, A. A., Testa, A. C., Veldman, J., Vergote, I., Van Huffel, S. Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery: prospective validation by IOTA group. BMJ, 2010. DOI: 10.1136/bmj.c6839

10. Alcázar JL, Pascual MÁ, Olartecoechea B, Graupera B, Aubá M, Ajossa S, Hereter L, Julve R, Gastón B, Peddes C, Sedda F, Piras A, Saba L, Guerriero S. IOTA simple rules for discriminating between benign and malignant adnexal masses: prospective external validation. Ultrasound Obstet Gynecol 2013. DOI: 10.1002/uog.12485.

11. Tantipalakorn C., Wanapirak C., Khunamornpong S., Sukpan K., Tongsong T. IOTA Simple Rules in Differentiating between Benign and Malignant Ovarian Tumors. Asian Pac J Cancer Prev 2014. DOI: 10.7314/APJCP.2014.15.13.5123

12. Koneczny J., Czekierdowski A., Florczak M., Poziemski P., Stachowicz N., Borowski D. 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. Ginekologia Polska 2017. DOI: 10.5603/GP.a2017.0116

13. Meys EM, Kaijser J, Kruitwagen RF, Slangen BF, Van Calster B, Aertgeerts B, Verbakel JY, Timmerman D, Van Gorp T. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis. Eur J Cancer 2016. DOI: 10.1016/j.ejca.2016.01.007

14. Kaijser J, Sayasneh A, Van Hoorde K, Ghaem-Maghami S, Bourne T, Timmerman D, Van Calster B. Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: a systematic review and meta-analysis. Hum Reprod Update 2014. DOI: 10.1093/humupd/dmt059.

15. Auekitrungrueng, Rapeepat, Tinnangwattana, Dangcheewan, Tantipalakorn, Charuwan et al. Comparison of the diagnostic accuracy of International Ovarian Tumor Analysis simple rules and the risk of malignancy index to discriminate between benign and malignant adnexal masses. International Journal of Gynecology & Obstetrics 2019. DOI: 10.1002/ijgo.12891

16. He RH, Yao WM, Wu LY, Mao YY. Highly elevated serum CA-125 levels in patients with non-malignant gynecological diseases. Arch Gynecol Obstet 2011. DOI: 10.1007/s00404-010-1717-5

17. Arun-Muthuvel V., Jaya V. Pre-operative Evaluation of Ovarian Tumors By Risk of Malignancy índex, CA-125 and Ultrasound. Asian Pacific Journal of Cancer Prevention 2014. Doi: 10.7314/APJCP.2014.15.6.2929

18. Froyman, Wouter, Timmerman, Dirk. Methods of Assessing Ovarian Masses. Obstetrics and Gynecology Clinics of North America 2019. DOI: 10.1016/j.ogc.2019.07.003

19. Valentin L, Jurkovic D, Van Calster B, Testa A, Van Holsbeke C, Bourne T, et al. Adding a single CA 125 measurement to ultra- sound imaging performed by an experienced examiner does not improve preoperative discrimination between benign and malignant adnexal masses. Ultrasound Obstet Gynecol 2009. Doi: 10.1002/uog.6415

**ENDEREÇO PARA CORRESPONDÊNCIA**
Ângela Rodrigues
Centro Hospitalar e Universitário de Coimbra EPE
Coimbra, Portugal.
E-mail: angela.reisrodrigues@gmail.com

**RECEBIDO EM:** 10/04/2020
**ACEITE PARA PUBLICAÇÃO:** 09/06/2020