# Comparative study on the diagnostic accuracy of the different international ovarian tumor analysis (IOTA) predictive model in discriminating between benign and malignant ovarian new growths: Logistic regression 1 and 2 (LR1 and LR2) and assessment of the different neoplasias of the adnexa (ADNEX) model*

By Marnie Ann Espiritu-Concepcion, MD, DPOGS and Kareen N. Reforma, MD, FPOGS, FPSUOG
Department of Obstetrics and Gynecology, Philippine General Hospital, University of the Philippines-Manila

**ABSTRACT**

**Objective:** To compare the diagnostic accuracy of the International Ovarian Tumor Analysis (IOTA)- Logistic Regression 1 and 2 (LR1 and LR2) and the Assessment of Different Neoplasias in the Adnexa (ADNEX) model in discriminating between benign and malignant ovarian new growths

**Methods:** The study was a prospective validation study. It included all patients admitted at the Gynecology ward of the Philippine General Hospital for elective surgery for ovarian new growths. Demographic information and clinical data were recorded for eligible patients. Two-dimensional ultrasonography with Doppler studies were performed. Ovarian new growths were classified based on IOTA LR1, LR2 and ADNEX model. Correlation of the ultrasound findings with the histopathology report and final staging based on Federation of Gynecology and Obstetrics (FIGO) classification was done.

**Results:** Sixty seven (67) patients were included in the final analysis. The mean age was 43 years old (range of 17-78). There were sixteen (16) nulligravid patients (22%). Eighteen (18) out of the 67 patients (27%) had malignant ovarian masses on histopathology. The IOTA LR1 had an area under the curve (AUC) of 0.96, sensitivity of 89% (95%CI, 74-100) and specificity of 92% (95%CI, 84-100). The IOTA LR2 had an AUC of 0.88, sensitivity of 61% (95%CI, 39-84) and specificity of 96% (95%CI, 90-100). The IOTA ADNEX had an AUC of 0.96, sensitivity of 89% (95%CI, 74-100) and specificity of 76% (95%CI, 63-88). Sensitivity and specificity of IOTA ADNEX for the diagnosis of specific malignant subtypes were as follows: Borderline, 80% and 76%, Stage I, 100% and 100, Stage II-IV, 86% and 100%. Accuracy values were not computed for the metastatic cancer since there was only one case seen. There was no significant difference in the accuracy values of IOTA ADNEX with or without CA 125.

**Conclusion:** In conclusion, IOTA LR1, LR 2 and ADNEX models were all useful tools in discriminating between benign and malignant ovarian masses. IOTA LR1 had the highest accuracy in differentiating between benign and malignant ovarian masses.

*Keywords: Ovarian new growth, Ultrasonography with Doppler studies, IOTA LR1, LR2 and ADNEX model*

## INTRODUCTION

Ovarian new growth is a common gynecologic diagnosis. It may present as a benign cystic mass or a malignant solid tumor. Premenopausal women usually have benign masses. The overall incidence of malignancy of an ovarian new growth in premenopausal women is 1:1000 increasing to 3:1000 at the age of 50.[1] Ovarian malignancy is the seventh most common cause of cancer in women world- wide and fifth in the Philippines. It is the eight leading cause of death in women worldwide.[2,3] It is an aggressive and fatal cancer, which is usually diagnosed at an advanced stage.

Based on the 2016 annual report of the Ultrasound section of the Department of Obstetrics and Gynecology at the Philippine General Hospital, twenty eight percent (3,083 out of 10,987) of the total gynecologic patients for the year were diagnosed to have ovarian new growths. While according to the 2016 surgical pathology statistics of the department, fourteen percent (518 out of 3,747) of the total admitted gynecological patients have ovarian masses and subsequently underwent surgery.

* First Place, Philippine Society of Ultrasound in Obstetrics and Gynecology. (PSUOG) Research Paper Contest, October 12, 2018, Satellite Office, Dr. Jose Fabella Memorial Hospital - Philippine Blood Disease and Transfusion Center, Quezon Avenue, Quezon City

Accurate preoperative discrimination between benign and malignant ovarian new growth is crucial in clinical practice. Patient's management will greatly depend on the pre-operative diagnosis. Benign masses are managed conservatively by laparoscopy and fertility sparing surgery while malignant lesions have extensive surgery and undergo complete surgical staging. A thorough clinical history and complete physical examination with ultrasonography with Doppler studies, particularly done by an expert sonographer, can precisely diagnose a benign and malignant ovarian new growth. However, an experienced sonographer is not always available.[4]

Several ultrasound-based prediction models and scoring systems have been created to aid the less experienced sonographers in differentiating between benign and malignant ovarian masses.[4] Sassone scoring system established in 1991 and 1Lerner scoring system established in 1994, were the initial methods used. To further increase the sensitivity and specificity of ultrasonography, the IOTA group developed various predictive models to calculate the risk of malignancy. It created the simple rules and logistic regression, LR1 and LR2, which are mathematical risk models that discriminate between benign and malignant ovarian masses. Its most recent publication is the IOTA ADNEX model, which is the first predictive multi-class model.[5] It can classify ovarian new growth between benign, borderline, stage I, stage II-IV and secondary metastatic cancer.[5] The reference standard for the different IOTA predictive models is the histopathological report of the excised ovary and surgical staging using the International Federation of Gynecology and Obstetrics (FIGO) classification for the malignant tumors.[5]

In addition to ultrasound with Doppler studies, serum CA-125 is requested for patients suspected to have ovarian malignancy. It is one of the parameters in the IOTA ADNEX model. It aids in the preoperative diagnosis but more importantly, it is used for monitoring the response to treatment.

The aim of this study to compare the diagnostic accuracy of the IOTA LR1, LR2 and ADNEX model in discriminating between benign and malignant ovarian new growths. The best predictive model can be utilized by the Ultrasound section of the Department of Obstetrics and Gynecology at the Philippine General Hospital as the standard model in discriminating ovarian new growths.

## OBJECTIVES

**General objective:**

To compare the diagnostic accuracy of IOTA LR1, LR2 and ADNEX model in discriminating between benign and malignant ovarian new growth

**Specific Objectives:**

1. To determine the clinical and demographic profile of the subjects diagnosed with benign and malignant ovarian masses.

2. To determine the sensitivity, specificity, likelihood ratios, positive and negative predictive values and area under the curve of IOTA LR1 model in discriminating between benign and malignant ovarian new growth.

3. To determine the sensitivity, specificity, likelihood ratios, positive and negative predictive values and area under the curve of IOTA LR2 model in discriminating between benign and malignant ovarian new growth

4. To determine the sensitivity, specificity, likelihood ratios, positive and negative predictive values and area under the curve of IOTA ADNEX model in discriminating between benign and malignant ovarian new growth

5. To compare the sensitivity, specificity, likelihood ratios, positive and negative predictive values and area under the curve of IOTA LR1, LR2 and ADNEX model in discriminating between benign and malignant ovarian new growth

## STUDY DESIGN

Prospective validation study

## MATERIALS AND METHODS

***Patient population:***

1. Study subjects

It included all patients admitted at the Gynecology ward of the Philippine General Hospital with sonographic finding of an ovarian new growth and subsequently underwent elective surgery from February 2018 to April 2018.

2. Inclusion Criteria:

Patients with sonographic finding of an ovarian new growth who underwent elective surgery within the period of February 2018 to April 2018.

3. Exclusion criteria:
   a. Patients with non gynecologic adnexal masses
   b. Pregnant at the time of diagnosis
   c. Patient with previous diagnosis and management of ovarian cancer at any stage
   d. Patient with ovarian masses for emergency surgery due to complication

***Sample Size***:

The sample size in this study was computed based on the findings from previous studies or reported statistics. The values reported from literature were utilized as parameters in the sample size equation appropriate for testing sensitivity and specificity.

Based on the initial study of Timmerman D., et al in 2005 on IOTA LR model, LR1 had a sensitivity of 93% and a specificity of 77% while LR2 had a sensitivity of 92% and a specificity of 75%. While the IOTA ADNEX yielded 96.5% sensitivity and 71.3% specificity, based on the study of Van Calster, B. et al in 2014 on IOTA ADNEX. The cut off for the total risk of malignancy was 10%.

Assuming that LR1, LR2, and IOTA ADNEX model would generate an average accuracy such 93.8% sensitivity and 74.4% specificity in detecting malignancy risk which occurred among the new cases of ovarian growths that occurred at a rate of 28% as reported in 2016 annual statistics of the Ultrasound section of the Department of Obstetrics and Gynecology at the Philippine General Hospital i.e. 3,083 out of 10,987 total gynecologic patients. Then using the sample size equation (see below), the needed sample size for this study was 65 cases estimated at +/-11.4% sampling error.

| | | |
|---|---|---|
| **Expected Sensitivity** | **0.94** | ◄ From literature or pilot study |
| **Expected Specificity** | **0.74** | ◄ From literature or pilot study |
| **Expected Prevalence** | **0.28** | ◄ From literature or pilot study |
| **Desired Precision** | **0.11** | ◄ Researcher's judgment |
| **Confidence level** | **95%** | ◄ 95% is recommended. |

Sample Size

**Explanation on the calculation of sample size in Sensitivity & Specificity Studies:**

| | | Disease + | Disease - |
|---|---|---|---|
| Test | + | *a* | *b* |
| | - | *c* | *d* |
| | | *(a+c)* | *(b+d)* |

*N*

**The sample size that we want to calculate is "total sample size" *N*.**

**What the study would like to determine are "Sensitivity" which is *a*/(*a*+*c*), and "Specificity" which is *d*/(*b*+*d*).**
**Therefore, we need to calculate the sample size to acquire an appropriate precision for estimating "Sensitivity" and "Specificity".**

**By using usual single proportion sample size formula (1), we can calculate the sample size for (*a*+*c*), if we use 'Sensitivity' as *P*. (Δ is precision).**
**After we get (*a*+*c*), the total sample size can be obtained by using Prevalence of the disease (using the formula 2).**
**Similary, we can calculate the sample size for (*b*+*d*), if we use specificy as *P* in formula (1). After getting (*b*+*d*), the total sample size can be calculated by formula (3).**

**FORMULA**

$$n = \frac{Z^2 * P(1-P)}{\Delta^2} \quad (1)$$

***n* will be (*a*+*c*) if we use Sensitivity as *P*, and *n* will be (*b*+*d*) if we use Specificity as *P* in formula (1).**

$$N = \frac{(a+c)}{\text{Prevalence}} \quad (2)$$

$$N = \frac{(b+d)}{(1-\text{Prevalence})} \quad (3)$$

***Description of the study procedure:***

All preoperative patients diagnosed with ovarian new growth and admitted for elective surgery were invited to participate in the study. The study procedure was fully explained to the patients and informed consent was obtained by the primary investigator. For minor participants, a parent or legally authorized representative (LAR) provided the necessary consent for the participation of the minor. Aside from this informed consent, assent from minors was also be obtained. No formal assent was needed for minors under 7 years old. A verbal assent was acceptable for participants 7 to under 12 years old. For minors 12 to under 15 years old, the participant signed a simplified assent form while for minors 15 to under 18 years old, the participant signed the similar informed document signed by the parents.

Patient data sheet containing demographic information and the value of the baseline serum Ca-125 for patients suspected to have ovarian malignancy, was accomplished by the primary investigator.

Samsung Accuvix A30 ultrasound model was utilized in the study. Transvaginal or transrectal ultrasound using 5-8 MHz transducer was performed by the primary investigator supervised by the co-investigator, within 24 to 48 hours prior to surgery. The patient with an empty bladder was placed in dorsal lithotomy position. The ultrasound probe was then inserted per vagina or rectum. For large masses, transabdominal ultrasound using 3-5 MHz transducer was performed while patient was in supine position. The entire duration of the procedure was approximately 30 minutes. In patients with bilateral ovarian masses with similar morphology, the larger mass was included in the study. For bilateral ovarian masses with different morphology, the risk of malignancy for each ovarian mass was separately calculated and included.

The ovarian new growth was classified as to whether it was benign or malignant based on the clinical and ultrasound parameters of the three IOTA models- LR1, LR2 and ADNEX.

The 10% cut off was used to predict the risk of malignancy. Furthermore, the IOTA ADNEX model differentiated among the 4 subgroups of malignancy namely borderline, stage I, Stage II-IV and secondary metastatic cancer. An absolute risk estimate, expressed in percentage was the outcome of this model. The results was then compared with the histopathologic diagnosis and final staging based on FIGO classification.

***Description of outcome measurements:***

The main outcome of this study was the comparison of the accuracy of IOTA LR1, LR2 and ADNEX model in discriminating between benign and malignant ovarian

new growths. This was evaluated based on the sensitivity, specificity, likelihood ratio, negative and positive predictive values and area under the curve of each IOTA model.

***Analysis of Data:***

All categorical data such as the family and personal history of ovarian cancer, current use of hormone therapy, personal and family history of breast cancer were expressed in frequency and percentages while continuous variable data such as age, gravity and parity were described by means and standard deviation. In testing association between LRI 1 cut off (below and above the cut off), LR2, and Adnex Model with Gold standard, 2x2 Fischer Exact test was used.

Moreover, Area Under Curve Index was plotted wherein an AUC of 0.70 and above was considered clinically acceptable accuracy performance. Moreover, the recommended cut off values of LR1 and LR2 as plotted by the AUCs were subjected to accuracy computation such as sensitivity, specificity, likelihood ratio, negative predictive and positive predictive values. IBMSPSS ver 21 was used as software.

## RESULTS

Between February to April 2018, a total of 73 patients with sonographic findings of an ovarian new growth, admitted for elective surgery were included in the study. Out of the 73 patients, 6 patients were excluded: One (1) patient expired prior to surgery, three (3) patients with concomitant medical diseases were not cleared for surgery and two (2) patients with final histopathology of paratubal cysts. Thus, a total of 67 patients were included in the final data analysis.

The overall mean age of the the patient was 43 years old (range of 17-78 years old). There were 16 (22%) nulligravid patients. The number of patients with current hormone therapy was 21 (31.3%). There was 1 (1.5%) patient with family history of ovarian cancer while there were 3 (4.5%) with family history of breast cancer.

The final histopathology results revealed 49 (73.1%) benign, 5 (7.5%) borderline and 13 (19.4%) malignant ovarian masses. Among the malignant masses, 3 (23.1%) patients were classified as stage I, 9 (69.2%) as stage II-IV and 1 (7.7%) as secondary metastatic cancer. The 3 most common benign ovarian masses (Table 1) were mature cystic teratomas (n=15; 22.4%), mucinous cystadenomas (n= 13; 19.4%) and endometriotic cysts (n=10; 14.9%). The most common malignant ovarian tutors (Table 1) were granulosa cell tumors (n= 3; 4.5%), mucinous cystadenocarcinomas (n=2 ; 3.0%) and mixed carcinomas (n=2; 3.0%).

There were two cases classified as malignant in all 3 IOTA models however, final histopathology revealed serous cystadenofibroma and mature cystic teratoma with endometriosis cyst. On the other hand, 2 cases were diagnosed benign in all 3 IOTA models but, on final histopathology showed squamous cell and mucinous carcinoma, arising from mature teratoma and serous borderline tumor.

**Table 1.** Pathology Results of 67 Ovarian Masses

| PATHOLOGY | n (%) |
|---|---|
| **Benign** | **49 (73.1%)** |
| Mature Cystic Teratoma | 15 (22.4%) |
| Mucinous Cystadenoma | 13 (19.4%) |
| Endometriotic Cysts | 10 (14.9%) |
| Serous Cystadenoma | 4 (6.0%) |
| Seromucinous Cystadenoma | 3 (4.5%) |
| Mixed Benign Cyst | 2 (3.0%) |
| Serous Cystadenofibroma | 1 (1.5%) |
| Benign Cyst | 1 (1.5%) |
| **Borderline** | **5 (7.5%)** |
| Serous | 3 (3.0%) |
| Mucinous cystadenoma | 2 (4.5%) |
| **Malignant** | **13 (19.4%)** |
| Granulosa Cell Tumor | 3 (4.5%) |
| Mucinous Cystadenocarcinoma | 2 (3.0%) |
| Mixed Carcinoma | 2 (3.0%) |
| Clear Cell Carcinoma | 1 ( 1.5%) |
| Transitional Cell Carcinoma | 1 (1.5%) |
| Endometrioid Adenocarcinoma | 1 (1.5%) |
| Metastatic Signet Ring Carcinoma | 1 (1.5%) |
| Adenocarcinoma | 1 (1.5%) |
| Papillary Cystadenocarcinoma | 1 (1.5%) |

### VALIDATION OF IOTA LR1

The IOTA LR1, at a cut-off of more than or equal to 10%, had a sensitivity (Table 3) of 89% (95%CI, 74-100) and a specificity of 92% (95%CI, 84-100). The AUC (Figure 1) for the overall discrimination between benign and malignant ovarian masses was 0.96. The sensitivity was lower than the initial IOTA LR study done by Timmerman, D. et.al (2005), which was 93%, while the specificity was higher than the initial IOTA LR study, which was 77%.

### VALIDATION OF IOTA LR2

The IOTA LR2, at a cut-off of more than or equal to 10%, had a sensitivity (Table 3) of 61% (95%CI, 39-84) and a specificity of 96% (95%CI, 90-100). The AUC (Figure 1)for the overall discrimination between benign and malignant ovarian masses was 0.88. The sensitivity was lower than the initial IOTA study done by Timmerman, D. et.al (2005), which was 92% while the specificity was higher than the initial IOTA LR study, which was 75%.

**Table 2.** Distribution of Benign and Malignant Cases Based on Different Method of Assessments such as IOTA LR1, IOTA LR 2, IOTA ADNEX, IOTA ADNEX WITHOUT CA 125, and Histopathology (Gold Standard)

| Methods | HISTOPATH (Gold Standard) | | Total |
|---|---|---|---|
| | malignant | benign | |
| No. of Samples | 18 | 49 | 67 |
| IOTA LR1 | | | |
| Malignant | 16 | 4 | 20 |
| Benign | 2 | 45 | 47 |
| IOTA LR2 | | | |
| Malignant | 11 | 2 | 13 |
| Benign | 7 | 47 | 54 |
| IOTA ADNEX | | | |
| Malignant | 16 | 12 | 28 |
| Benign | 2 | 37 | 39 |
| IOTA ADNEX WITHOUT CA 125 | | | |
| Malignant | 16 | 11 | 27 |
| Benign | 2 | 38 | 40 |

## VALIDATION OF IOTA ADNEX

The IOTA ADNEX, at a cut-off of more than or equal to 10%, had a sensitivity (Table 3) of 89% (95%CI, 74-100) and a specificity of 76% (95%CI, 63-88). The AUC (Figure1) for the overall discrimination between benign and malignant ovarian masses was 0.96. The sensitivity was lower than the initial IOTA ADNEX study by Van Calster, B, et.al (2015), which was 96.5% while the specificity was higher than the initial IOTA ADNEX study, which was 71.3%.

The sensitivity (89%) was the same even without the CA-125. But the specificity of IOTA ADNEX with CA-125 (76%) was slightly higher in IOTA without the CA-125 (78%). The AUC of IOTA ADNEX with CA-125 (0.96) was slightly lower than IOTA ADNEX without CA-125 (0.94) however, it is not statistically significant.

Sensitivity and specificity of IOTA ADNEX (Table 5) for the diagnosis of specific malignant subtypes were as follows: Borderline, 80% and 76%, Stage I: 100% and 100%, Stage II-IV, 86% and 100%. Accuracy values were

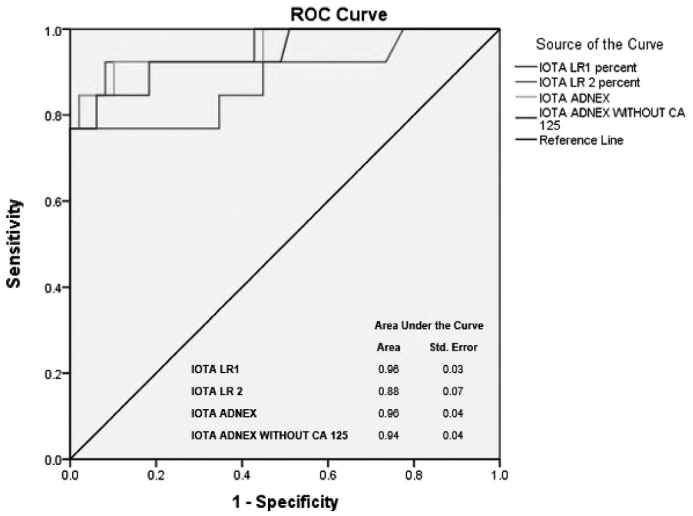


**Figure 1.** Area Under Receiver's Operating Curve (AUROC) of the Accuracies of IOTA LR1, IOTA LR 2, IOTA ADNEX and IOTA ADNEX WITHOUT CA 125 in Predicting Malignancy using the Histopathology as Gold Standard

**Table 3.** Accuracy of IOTA LR1, IOTA LR 2, IOTA ADNEX and IOTA ADNEX WITHOUT CA 125 in Predicting Malignancy using Histopathology as Gold Standard

| Accuracy Indices | IOTA LR1 | IOTA LR 2 | IOTA ADNEX | IOTA ADNEX WITHOUT CA 125 |
|---|---|---|---|---|
| | Value, CI 95% | Value, CI 95% | Value, CI 95% | Value, CI 95% |
| Sensitivity | 89 (74-100) | 61 (39-84) | 89 (74-100) | 89 (74-100) |
| Specificity | 92 (84-100) | 96 (90-100) | 76 (63-88) | 78 (66-89) |
| Likelihood Ratio + | 10.89 (4.2-28.25) | 14.97 (3.67-61.11) | 3.63 (2.16-6.09) | 3.96 (2.29-6.83) |
| Pred value positive | 80 (62-98) | 85 (65-100) | 57 (39-75) | 59 (41-78) |
| Pred value negative | 96 (90-100) | 87 (78-96) | 95 (88-100) | 95 (88-100) |

not computed for the metastatic cancer, since there was only one case seen.

Sensitivity and specificity of IOTA ADNEX without CA-125 (Table 7) for the diagnosis of specific malignant subtypes were as follows: Borderline, 80% and 78%, Stage I: 100% and 100%, Stage II-IV, 100% and 100%. Accuracy values were not computed for the metastatic cancer since there was only one case seen.

**COMPARISON OF IOTA LR1, LR2 AND ADNEX MODEL**

The IOTA LR1 and ADNEX had similar sensitivity which was higher than that of IOTA LR2. IOTA LR2, however, showed the highest specificity among the 3 predictive models. When comparing the overall performance in discriminating between benign and malignant ovarian masses, IOTA LR1 and ADNEX performed better than IOTA LR2, as expressed in their respective AUCs. Both IOTA LR1 and ADNEX had higher sensitivity than LR 2 but with lower specificity.

**Table 4.** Distribution of Benign and the Stages of Malignant Cases based on IOTA ADNEX and Histopathology (Gold Standard)

| IOTA ADNEX | Histopathology | | | | | |
|---|---|---|---|---|---|---|
| | Benign | Borderline | Stage I | Stage II-IV | Metastatic | Total |
| Benign | 37 | 1 | 0 | 1 | 0 | 38 |
| Borderline | 12 | 4 | 1 | 1 | 0 | 18 |
| Stage I | 0 | 0 | 2 | 1 | 0 | 14 |
| Stage II-IV | 0 | 0 | 0 | 6 | 1 | 7 |
| Metastatic | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 49 | 5 | 3 | 9 | 1 | 67 |

**Table 5.** Accuracy of IOTA ADNEX in discriminating among the stages of malignancy using histopathology as Gold Standard

| Accuracy | Histopathology | | | |
|---|---|---|---|---|
| | Bordeline | Stage 1 | Stage II-IV | Metastatic |
| | Value, CI 95% | Value, CI 95% | Value, CI 95% | Value, CI 95% |
| Sensitivity | 80 (45-100) | 100 (100-100) | 86 (60-100) | no estimate |
| Specificity | 76 (63-88) | 100 (100-100) | 100 (100-100) | no estimate |
| Likelihood Ratio + | 3.27 (1.69-6.31) | no estimate | no estimate | no estimate |
| Pred value positive | 25 (4-46) | 100 (100-100) | 100 (100-100) | no estimate |
| Pred value negative | 97 (92-100) | 100 (100-100) | 97 (92-100) | no estimate |

**Table 6.** Distribution of Benign and the Stages of Malignant Cases based on IOTA ADNEX without CA-125 and Histopathology (Gold Standard)

| IOTA ADNEX without CA-125 | Histopathology | | | | | |
|---|---|---|---|---|---|---|
| | Benign | Borderlin e | Stage I | Stage II-IV | Metastatic | Total |
| Benign | 38 | 1 | 0 | 0 | 0 | 38 |
| Borderline | 11 | 4 | 1 | 2 | 0 | 18 |
| Stage I | 0 | 0 | 1 | 0 | 0 | 4 |
| Stage II-IV | 0 | 0 | 1 | 6 | 1 | 7 |
| Metastatic | 0 | 0 | 0 | 0 | 0 | 0 |
| Total | 49 | 5 | 3 | 9 | 1 | 67 |

**Table 7.** Accuracy of IOTA ADNEX without CA-125 in discriminating among the stages of malignancy using histopathology as Gold Standard

| Accuracy | Histopathology | | | |
|---|---|---|---|---|
| | Bordeline | Stage 1 | Stage II-IV | Metastatic |
| | Value, CI 95% | Value, CI 95% | Value, CI 95% | Value, CI 95% |
| Sensitivity | 80 (45-100) | 100 (100-100) | 100 (100-100) | no estimate |
| Specificity | 78 (66-89) | 100 (100-100) | 100 (100-100) | no estimate |
| Likelihood Ratio + | 3.56 (1.8-7.04) | no estimate | no estimate | no estimate |
| Pred value positive | 27 (4-49) | 100 (100-100) | 100 (100-100) | no estimate |
| Pred value negative | 97 (92-100) | 100 (100-100) | 100 (100-100) | no estimate |

## DISCUSSION

Based on the study, all 3 IOTA models were useful tools in discriminating between benign and malignant ovarian masses as manifested in their AUCs above the cut off of 0.70. However, IOTA LR1 and ADNEX had better AUC of 0.94 than IOTA LR2 with AUC of 0.88.

IOTA LR 1 had similar sensitivity with ADNEX but with higher specificity. This can be attributed to certain parameters utilized in ADNEX but not in LR1, such as the number of locules and serum Ca 125. Five (5) ovarian masses were classified borderline malignant using IOTA ADNEX but benign using LR1. Three (3) of these ovarian masses had elevated CA-125 and 2 had more than 10 locules, thus, classifying them under borderline malignant tumors. On histopathology, all 5 cases were benign, with histologic diagnosis consistent with endometriotic cyst and mucinous cystadenoma.

IOTA ADNEX had more false positive results than LR1, thus resulting in lower specificity.

IOTA LR2 had the lowest sensitivity but the highest specificity. This may be due to limited parameters of IOTA LR2. It did not include the largest diameter of the ovarian mass which was included in both LR 1 and ADNEX models, as well as the number of locules, evaluated in the ADNEX model. In IOTA LR2, an ovarian mass had a high probability of malignancy if there were solid areas with detectable blood flow and ascites. In IOTA LR1 and ADNEX, large multiloculated ovarian masses even without solid areas and ascites were classified as malignant, particularly borderline malignant in IOTA ADNEX.

In this study, IOTA LR 1 and ADNEX had lower sensitivity but higher specificity compared with the initial IOTA LR by Timmerman, D. et.al (2005) and ADNEX study by Van Calster, B et.al (2014). The lower sensitivity may be attributed to smaller sample size of this study compared with the original study. There were actually only 2 false negative results based on IOTA LR1and ADNEX.

IOTA LR2 also had a significantly lower sensitivity and higher specificity than the initial IOTA LR study by Timmerman, D. et.al (2005). This can be ascribed to the difference in the study population. The study was conducted in a tertiary government hospital with a gynecologic oncology center, which was referral center from the different provinces of the Philippines. Thus, most patients seen already had huge ovarian masses. Since, the largest diameter was not one of the parameters of LR2, these masses were classified as benign, which most of them turned out to be borderline malignant on histopathology.

Aside from discriminating between benign and malignant ovarian masses, IOTA ADNEX model calculates the probability that a malignant ovarian mass is borderline malignant, stage I, stage II-IV and secondary metastatic cancer. In classifying the borderline malignant tumors, IOTA ADNEX had a good sensitivity of 80% but a low specificity of 76%. Ovarian masses classified as borderline malignant in IOTA ADNEX were large multiloculated (>10 locules) masses.

There were 5 borderline malignant ovarian masses diagnosed on histopathology. Four (4) out of 5 were accurately diagnosed. These were large multiloculated (>10 locules) ovarian masses ranging from 175 mm to 362 mm. Two (2) of them had elevated serum CA-125. The borderline malignant ovarian tumor diagnosed as benign by IOTA ADNEX model was a unilocular cyst measuring 110 mm with normal serum CA 125. The sensitivity and specificity in classifying stage 1 and stage II-IV were relatively high. The accuracy values for secondary metastatic cancer were not obtained because there was only one case. The percentage of malignant cases (27%) in the study was at par with the previous IOTA ADNEX study by Van Calster, B. et al. (2015), however, the sample size was smaller. This limited the sub-analysis of the accuracy IOTA ADNEX in classifying the ovarian masses among the different subgroups of malignancy.

All patients in this study had serum CA-125. There was no significant difference in the diagnostic accuracy values of IOTA ADNEX with or without CA 125 in discriminating between benign and malignant ovarian masses and in differentiating the different malignant subgroups.

## CONCLUSION

In conclusion, IOTA LR1, LR 2 and ADNEX models were all useful tools in discriminating between benign and malignant ovarian masses. However, IOTA LR1 and ADNEX were better than IOTA LR2, as manifested in their AUCs. IOTA LR1 and ADNEX had similar sensitivities but LR1 had a higher specificity. Therefore in this study, IOTA LR1 had the highest accuracy in differentiating between benign and malignant ovarian masses.

IOTA ADNEX had an added advantage of being polytomous model for the differentiation between the various subtypes of malignancy, however in the study, the sub -analysis of its accuracy was limited by small sample size of the malignant cases.

## RECOMMENDATION

Although the study was able to reach its target sample size, the small percentage of malignant ovarian masses limited the sub-analysis of the accuracy of IOTA ADNEX in differentiating among the different subtypes of malignancy. It is therefore recommended to further increase the sample size to better assess the accuracy of IOTA ADNEX in differentiating between the different subgroup of malignancy.

Since IOTA LR1 was noted to have highest accuracy, it can be utilized by the section of Ultrasound in discriminating between benign and malignant ovarian masses. ■

## REFERENCES


1. Royal College of Obstetrician and Gynecologists and British Society Gynecological Endoscopy Joint Guideline. (2011). Management of Suspected Ovarian Masses in Premenopausal Women.
2. World Cancer Research Fund / American Institute for Cancer Research Continuous Update Project Report. Food, Nutrition, Physical Activity, and the Prevention of Ovarian Cancer 2014. http://www.dietandcancerreport.org/cup/cup_resources.php.
3. Department of Health: Ovarian Cancer. http://www.doh.gov.ph
4. Mets, E.M.J. et al. (2017). Estimating risk of malignancy in adnexal masses: external. validation of the ADNEX model and comparison with other frequently used ultrasound methods. *Ultrasound Obstetrics and Gynecology*; 49, 784-792.
5. Van Calster, B. et al. (2015). Practical guidance for applying ADNEX model from the IOTA group to discriminate between the different subtypes of adnexal tumors. Facts Views Vis OBGYN, 7, 32-41.
6. Timmerman, D. et al. (2000). Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) group. *Ultrasound Obstetrics and Gynecology*; 16, 500-505.
7. International Oavarian Tumor Analysis 2017. http://www.iotagroup.org
8. Timmerman, D. et al. (2005). Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multi center study by the International Ovarian Tumor Analysis group. *Journal of Clinical Oncology*, 23:8794-8801.
9. Kaijser, J. et. Al. (2015). Towards an evidence-based approach for the diagnosis and management of adnexal masses: findings of the International Ovarian Tumor Analysis studies. *Facts Views Vis Obgyn*, 7(1):42-59.
10. Kaiser, J et.al. (2014). Presurgical diagnosis of adnexal tumors using mathematical models and scoring systems: a systematic review and meta- analysis. *Hum Reproduct Update*, 20:449-62.
11. Nunes, N. et al. (2013). A prospective validation of the IOTA logistic regression models (LR1 and LR2) in comparison to subjective pattern recognition for the diagnosis of ovarian cancer. *Int J Gynecol Cancer*, 23:1583-9.
12. Bolo-Paiso, A. and Reforma, K. (2017). The Diagnostic Accuracy of International Tumor Analysis (IOTA) - Logistic Regression 2 in Tertiary Government Hospital. Unpublished manuscript.
13. Van Calster, B. et al. (2014). 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 tumors: prospective multi center diagnostic study. *BMJ*, 349: g5920.
14. De Vera, M.U. (2016) Diagnostic Accuracy of IOTA-ADNEX model in differentiating, benign, borderline, early and advanced stage invasive and secondary metastatic adnexal masses: a global South experience. Ultrasound in Obstetrics and Gynecology, Book of Abstracts, 48: p201.
15. Sayasneh, A. et al. (2013). A multicenter external validation of IOTA prediction models and RMI by operators with varied training. *Br J Cancer*, 108:2448-54.
16. Van Holsbke C. et al. (2012). External validation of diagnostic models to estimate the risk of malignancy in adnexal masses. *Clin Cancer res*, 18:815-25.
17. Ultrasound section-Department of Obstetrics and Gynecology of the Philippine General Hospital (2017). Ultrasound section 2016 Annual Report. Chia, C.
18. Department of Obstetrics and Gynecology-Philippine General Hospital (2017). 2016 Annual Surgical Pathology Statistics. Fajuatagana et. al