Comparison of IOTA three-step strategy and logistic regression model LR2 for discriminating between benign and malignant adnexal masses

Juan José Hidalgo¹², Antoni Llueca³, Irene Zolfaroli², Nadia Veiga⁴, Ester Ortiz⁴, Juan Luis Alcázar⁵

¹Department of Obstetrics and Gynaecology, Hospital Comarcal Universitario de Vinaros, Castellón, ²Department of Obstetrics and Gynaecology, Hospital Clínico Universitario de Valencia, ³Department of Medicine. University Jaume I, Castellón, ⁴Department of Obstetrics and Gynaecology, Hospital Peset, Valencia, ⁵Department of Obstetrics and Gynaecology, Clinica Universidad de Navarra, University of Navarra, Pamplona, Spain

Introduction

An accurate initial diagnosis of benign or malignant ovarian masses is essential to guide the treatment [1]. Lesions with suspected malignancy must be referred to specialized gynaecological oncology centres because the correct initial surgical staging and cytoreduction surgery are some of the most important prognostic factors [2-6]. Subjective ultrasound assessment by experts has proven to be the most accurate method for characterizing adnexal masses [7-9], with sensitivity of 88-98% and specificity of 89-96% to establish their malignancy probability [10,11]. Nevertheless, it presents some limitations related to the subjectivity, experience and skill of the sonographer, showing limited diagnostic accuracy in less experienced professionals [12]. This is the reason why objective predictive models have been proposed to

Abstract

**Aims:** To compare the diagnostic performance of two ultrasound-based diagnostic systems for the classification of benign or malignant adnexal masses, the three-step strategy and the predictive logistic regression model LR2, both proposed by the International Ovarian Tumour Analysis (IOTA) Group. **Material and methods:** Prospective observational study at a single centre that included patients diagnosed with a persistent adnexal mass by transvaginal ultrasound over a period of two years. They were evaluated by a non-expert sonographer by applying the three-step diagnostic strategy and the LR2 predictive model to classify the masses as benign or malignant. Patients were treated surgically or followed up for at least one year, taking as the standard reference for benignity or malignancy the histological diagnosis of the lesion or ultrasound changes suggestive of malignancy during the follow-up period. Sensitivity, specificity, positive and negative likelihood ratios and overall accuracy of both systems was calculated and compared. **Results:** One hundred patients were included, with a mean age of 50.6 years (range 18-87). Surgery was performed on 62 (62%) patients and 38 (38%) were managed expectantly. Eighty-three (83%) lesions were benign and 17 (17%) were malignant. The IOTA three-step strategy presented sensitivity of 94.1% (95%CI, 86.7-98.3%) and specificity 97.6% (95%CI, 94.8-99%). The LR2 logistic regression model showed sensitivity 94.1% (95%CI, 73-98.9%) and specificity 81.9% (95%CI 72.3-88.7%). Comparison of the two systems showed a statistically significant difference in specificity in favour of the three-step strategy. **Conclusions:** The IOTA three-step strategy, in addition to being simple to use in clinical practice, has a high diagnostic accuracy for the classification of benignity and malignancy of the adnexal masses, overtaking that of other predictive models such as the LR2 logistic regression model.

Keywords: adnexal masses; diagnosis; ultrasound; three-step strategy; LR2
help non-expert examiners to achieve similar results to experts [13]. To date, none of them has been generally accepted in clinical practice nor have they been shown to improve the results obtained by the subjective assessment of experts [14,15].

One of these models, named clinically the oriented three-step strategy, was proposed by the International Ovarian Tumour Analysis (IOTA) Group in 2012, showing better results than previous models and a closer approach to clinical reality [16]. Five studies have been reported performing an external validation of this approach, showing a sensitivity and specificity of 87.5-95.2% and 87.6-100%, respectively [17-21]. With this strategy, adnexal lesions are evaluated in three consecutive steps. The first step evaluates six simple variables or descriptors (four for benignity and two for malignancy) of an immediate application. If none are applicable or both benign and malignant can be applied, the lesion is considered non-qualifiable and must be evaluated by a second step according to the Simple Rules. These first two steps can be performed by non-expert sonographers with basic training. If the diagnosis cannot be made by the second step, a third step should be applied consisting of subjective evaluation by an expert sonographer [16].

Another predictive system is the logistic regression model LR2, one of the most evaluated and most widely used adnexal lesion diagnostic systems in clinical practice, proposed by IOTA in 2005 [22]. This model estimates the individual probability of risk of malignancy of an adnexal mass applying to a logistic regression formula the result of six predictor variables. A probability result greater than 0.1 (10%) implies that the lesion should be considered malignant. Some studies have validated this model, showing that sensitivity for the diagnosis of malignancy is above 90% [13,22,23]. Furthermore, a systematic review and meta-analysis in 2014 found this to be one of the most accurate predictive models among the 19 evaluated [15].

To our knowledge, there are no published studies to date comparing the clinically oriented three-step strategy with other predictive models. In our study, our objective is to compare the diagnostic accuracy of the three-step strategy with that of the LR2 model in the same group of patients.

Materials and methods

Design of the study

An observational, prospective study was performed in a single centre (University District Hospital of Vinaros, Castellón, Spain) from September 2015 to August 2017.

Patients

Patients were eligible if they were over the age of 18 and were diagnosed with at least one adnexal mass by transvaginal or transrectal ultrasound. To remain in the study, they should have undergone surgery within three months after diagnosis or complete clinical and ultrasound follow-up for at least 12 months.

In the case of bilateral masses, the most complex or the largest (in case of morphological similarity) was selected. Patients who were pregnant during the study period were excluded. Likewise, patients who could not complete the follow-up visits or had a history or presence of a gynaecological neoplastic process were also excluded.

The Medical Ethics Committee of the centre approved the study. All patients received written and verbal information and signed their informed consent to participate. The patients included in the study constitute a subgroup within a more extensive study for the external validation of the three-step ultrasound strategy [21].

Ultrasound evaluation

The patients in the study were initially evaluated by a non-expert sonographer (J.J.H.), with a Level-2 training according to the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) for gynaecological ultrasound [24].

The ultrasound study was performed using transvaginal or transrectal B-mode ultrasound and 2D Power Doppler according to the IOTA terms, definitions and measurements [25]. Large lesions that could not be entirely evaluated by vaginal ultrasound were also evaluated abdominally. At least ten ultrasound images and two video fragments were electronically stored for each mass. The scans were performed with a Voluson S6® ultrasound system with RIC 5-9 MHz (GE Health care Ultrasound, Milwaukee, WI, USA).

Three-step strategy

At the initial assessment, the first step of the three-step IOTA strategy, consisting of the evaluation of six simple descriptors was applied (Table I) [16]. If the lesion could not be classified with this first step (none of the six descriptors was applicable or benignity and malignancy descriptors were present), it was evaluated with the Simple Rules, the second step of the system (Table I) [26].

Finally, if the mass could not be classified as benign or malignant with the Simple Rules (none of the features was present or they were both malignant and benign), the diagnosis was inconclusive, and their images were referred for subjective assessment by an expert sonographer (J.L.A.), with training equivalent to Level-3 of the EFSUMB [24].
Comparison of IOTA three-step strategy and logistic regression model LR2

Once the lesion was classified using the three-step strategy, it was evaluated using the LR2 model for calculating the malignancy probability assessing six variables: (1) age, (2) ascites, (3) papillary blood flow, (4) maximum solid component diameter, (5) irregular internal cystic walls and (6) acoustic shadows. The result was applied to the logistic regression formula $y = \frac{1}{1 + e^{-z}}$, where $z = -5.3718 + 0.0354 (1) + 1.6159 (2) + 1.1768 (3) + 0.0697 (4) + 0.9586 (5) - 2.9486 (6)$, using IOTA Models® application software, available with IOS® operating system. A probability above 10% was considered a high risk of malignancy [22].

**Therapeutic attitude**

The therapeutic attitude consisted of surgical treatment or expectant clinical and ultrasound follow-up. Patients who showed high risk of malignancy were addressed to the referral oncology gynaecology unit. Patients with masses considered benign symptomatic were referred to their general gynaecologist. In both cases, the therapeutic attitude was determined by their referring gynaecologist. Asymptomatic patients with a low-risk lesion were offered surgical treatment by a general gynaecologist or expectant management with clinical and ultrasound follow-up at 3, 6 and 12 months during the first year, and annually thereafter.

The referring gynecologist decided whether to determine serum tumor markers during the follow-up of the patients in whom a conservative attitude was chosen. In these cases, the value of these markers was not known by the investigating sonographer at the time of successive follow-up ultrasound evaluations.

**Reference standard**

Lesions managed expectantly were considered benign if they did not show ultrasound changes suggesting malignancy for at least one year. If the lesion disappeared spontaneously before completing one year of follow-up, the patient was excluded from the study. If any ultrasound scan revealed morphological changes in the lesion, the probability of malignancy was recalculated with the three-step strategy and the LR2 system. In the removed lesions, the standard reference was the histological diagnosis. Tumors that were histologically diagnosed as borderline were considered malignant for statistical analysis.

**Statistical analysis**

For the qualitative variables, the frequency distribution value for each of the categories was obtained, presenting the data in the form of absolute counts and percentages. Quantitative variables were studied following Kolmogorov-Smirnov analysis to determine if the distribution of values was normal. Quantitative variables with normal distribution were expressed as a central tendency measure with the standard deviation and range. When they did not follow a normal distribution, they were presented as median and range.

Sensitivity, specificity, and positive (LR+) and negative (LR-) likelihood ratio were calculated with a 95% confidence interval of the three-step model in order to discriminate between benign and malignant adnexal lesions. The sensitivity and specificity obtained with the application of the three-step model were compared with the McNemar test with those obtained with the logistic regression model LR2. A p<0.05 was considered statistically significant for all comparisons.

Statistical analysis was performed using the IBM SPSS version 20® for Windows (SPSS Inc., Chicago, IL).

**Results**

One hundred and two patients with an adnexal mass were studied. Two (1.9%) patients were excluded, one for a follow-up shorter than 12 months and the other for

| Simple descriptors [16] | Predictors of benignity | For predicting a benign tumour (B-rules) |
|-------------------------|--------------------------|-----------------------------------------|
|                         | • Unilocular tumour       | B1 Unilocular                           |
|                         |  in a premenopausal        | B2 Solid components where the largest   |
|                         |  woman                    |  has a largest diameter < 7 mm          |
|                         | • Unilocular tumour with   | B3 Acoustic shadows                      |
|                         |  mixed echogenicity       |                                        |
|                         |  and acoustic shadows in  |                                        |
|                         |  a premenopausal woman    |                                        |
|                         | • Unilocular anechoic      | B4 Smooth multilocular tumour with      |
|                         |  tumour with regular      |  largest diameter <100 mm               |
|                         |  walls and maximum       |                                        |
|                         |  diameter of lesion < 10  |                                        |
|                         |  cm                       |                                        |
|                         | • Remaining unilocular     | B5 No blood flow (color score 1)        |
|                         |  tumour with regular      |                                        |
|                         |  walls                    |                                        |

| Simple rules [26] | Predictors of Malignancy | For predicting a Malignant tumour (M-rules) |
|-------------------|--------------------------|---------------------------------------------|
|                   | • Tumour with ascites     | M1 Irregular solid tumour                   |
|                   |  and at least moderate   | M2 Ascites                                 |
|                   |  color Doppler blood      | M3 At least four papillary projections      |
|                   |  flow in a postmenopausal | M4 Irregular multilocular solid tumour      |
|                   |  woman                   |  with largest diameter ≥100 mm              |
|                   | • Age > 50 years and CA  | M5 Very strong blood flow (color score 4)   |
|                   | 125 > 100 U/mL           |                                            |

Table I. Simple descriptors (Step 1) and Simple Rules (Step 2) for classifying adnexal masses with IOTA three-step strategy.
being pregnant during the study. Therefore, 100 patients were included in the final analysis. The average age was 50.6 years (SD 17.4; range 18-87 years). Fifty-eight (58%) patients were premenopausal.

Sixty-two (62%) patients were surgically treated and the rest had follow-up for at least 12 months. Eighty-three (83%) lesions were benign. Out of the total surgically removed adnexal masses, 45 (72.6%) were benign and 17 (27.4%) malignant. The mean follow-up time was 14.2 months (SD 2.8; range 12-22 months). All patients in the expectant management group showed lesions that were considered benign at the end of the follow-up, showing no ultrasound or clinical changes suggestive of malignancy. Figure 1 shows a flow chart with the patient selection, their management and the final diagnostic result.

The mean age of patients with malignant masses was 51.6 (SD 17.7; range 22-85), whereas the mean age for women with benign lesions was 49.2 (SD 15.9; range 18-87). Benign lesions were diagnosed in 48 (57.8%) premenopausal and 35 (42.2%) postmenopausal patients. Malignant lesions occurred in 8 (47%) premenopausal and 9 (53%) postmenopausal women.

The most common histological diagnoses among the benign lesions removed were: 15 (33.3%) endometriomas and 12 (26.7%) serous cystadenomas. The most common malignant tumours were: 6 (35.3%) serous carcinomas and 3 (17.6%) mucinous carcinomas. Table II summarizes the histological diagnoses of all adnexal lesions removed with number of cases correctly and erroneously diagnosed in each step of the three-step strategy and the average LR2 outcome for each histological group.

Figure 2 shows a flowchart with the classification of benignity or malignancy of adnexal masses with the three-step strategy, indicating the diagnostic results of each of them according to the reference standard. With the first two steps, 91 (91%) lesions could be classified by non-expert sonographers (37 (37%) with simple descriptors (First step); SR: Simple Rules (Second step); FP: False Positive; FN: False Negative; §: One further case was diagnosed as malignant (false positive); ¥: One further case was diagnosed as benign (false negative).
Comparison of IOTA three-step strategy and logistic regression model LR2

scripts and 54 (54%) using Simple Rules. An expert sonographer classified the remaining 9 (9%) lesions. With the first step (simple descriptors) no malignant lesion was classified as benign, and only one benign was diagnosed as malignant (false positive). With the second step (Simple Rules), a malignant lesion was classified as benign (false negative) and a benign one as malignant (false positive). With the third step (expert ultrasound evaluation) all lesions were correctly classified. Figure 3 show the characteristics of the lesions that were misclassified by the three-step strategy.

The final diagnostic performance of this strategy for the whole study population was: sensitivity 94.1\% (95\% CI, 86.7-98.3\%); specificity 97.6\% (95\% CI, 94.8-99\%); LR+ 39.2 (95\% CI, 16.6-100.2); and LR-0.06 (95\% CI, 0.02-0.15). The diagnostic accuracy was 97\%.

With the application of LR2 taking a result above 10% as a high probability of malignancy, 31 (31\%) lesions were diagnosed as malignant and 69 (69\%) as benign. The mean probability of malignancy for all lesions was 18\% (SD 27.3; range 0.2-96.3). Of the 31 lesions classified as malignant, 16 (51.6\%) were definitely malignant and 15 (48.4\%) benign (false positives) according to the reference standard. Of the 69 lesions classified as benign, 68 (98.5\%) were benign and 1 (1.5\%) was malignant (false negative). The lesion that was false-negative with LR2 gave the same result when applying the three-step strategy.

The diagnostic performance of the LR2 model was: sensitivity 94.1\% (95\% CI, 73-98.9\%); specificity 81.9\% (95\% CI, 72.3-88.7\%); LR+ 5.2 (95\% CI, 3.2-8.3); and LR-0.07 (95\% CI, 0.01-0.48), being the diagnostic accuracy 84\%.

Comparison of the two diagnostic models using the McNemar test was statistically significant (p<0.05), which means that, given that the sensitivity for both was identical, the three-step model is more specific in the classification of lesions than LR2.

There was an agreement in the diagnosis of benignancy or malignancy between the two systems in 87 (87\%) patients. The 13 discordant cases (13\%) corresponded to lesions diagnosed as benign with the three-step strategy and malignant with LR2. In all cases, the final diagnosis according to the reference standard was benign, so all these cases corresponded to false positives from the LR2 model. Regarding the timing of the three-step strategy in
which each of the discordant cases was classified, nine were applied the Simple Rules and four the subjective assessment by the expert. Table III shows the characteristics of these discordant cases between the two diagnostic systems.

**Discussion**

Our study aims to compare the diagnostic accuracy of two ultrasound-based predictive systems, the three-step strategy and the logistic regression model LR2, both proposed by IOTA, in a group of patients with adnexal masses for classification as benign or malignant of those lesions. Our results show that the three-step strategy presents better diagnostic performance than the LR2 model and, therefore, the former should preferably be used in the appraisal of the adnexal masses by non-expert sonographers.

Both systems demonstrated similar sensitivity (94.1%) for the prediction of malignancy, which implies that their diagnostic capacity for patients with ovarian cancer is high. Therefore, both models would have a low rate of false negatives, meaning that few patients with a malignant pathology would be diagnosed as benign, an essential parameter in the diagnosis of ovarian cancer. However, we observed a statistically significant difference in specificity in favour of the three-step strategy (97.6% vs 81.9%), which would mean that fewer patients with benign lesions would be misdiagnosed with malignancy, avoiding unnecessary surgical interventions and their potential complications, as well as improving the economic costs and psychological repercussions in these patients.

In this sense, the results of LR+ and LR- also showed a higher performance of the three-step strategy compared to LR2 by presenting results of 39.2 and 0.06 respectively (vs 5.2 and 0.07 of LR2), which would indicate a better practical utility of the first system, allowing the confirmation of the presence of malignancy in the adnexal lesions with greater certainty, since a diagnostic test with LR+>10 and LR-<0.1 is highly relevant and useful [27].

So far, five external validation studies of the IOTA three-step strategy have been published. However, none of them has been compared with other predictive models in the same group of patients [17-21]. We chose the LR2 logistic regression system as a comparison model since it is one of the most evaluated, best-performing and most used adnexal lesion diagnostic systems in clinical practice [13,15,22,23]. This is why other diagnostic models have used this system as a reference in comparative studies, with the results being similar [28] or favourable to LR2 [23,29-31].

This would show that the three-step strategy has proven more diagnostic accuracy than one of the reference systems in the classification of adnexal lesions, that is, the LR2.

The main support of our study relies on its performance since it has been developed under conditions similar to the usual clinical practice. We included both surgically treated and expectantly managed lesions with clinical and ultrasound follow-up, a therapeutic attitude prevailing in many patients with adnexal masses. In addition, patients were evaluated by non-expert sonographers and in a non-referral centre for ultrasound diagnosis of ovarian cancer, also a common clinical practice for most

| Age | Three-step classification (Diagnostic step) | LR2 classification (Malignancy probability) | Histological diagnosis | Ultrasound follow-up (Diagnosis) |
|-----|------------------------------------------|---------------------------------------------|------------------------|---------------------------------|
| 1   | Benign (SR)                              | Malignant (17.8%)                          | Tuboovarian abscess    | -                               |
| 2   | Benign (SR)                              | Malignant (10.2%)                          | -                      | Yes (Benign)                    |
| 3   | Benign (EA)                              | Malignant (48%)                            | -                      | Yes (Benign)                    |
| 4   | Benign (SR)                              | Malignant (14.6%)                          | Serous cystadenoma     | -                               |
| 5   | Benign (SR)                              | Malignant (12.2%)                          | -                      | Yes (Benign)                    |
| 6   | Benign (SR)                              | Malignant (16.5%)                          | -                      | Yes (Benign)                    |
| 7   | Benign (SR)                              | Malignant (11%)                            | -                      | Yes (Benign)                    |
| 8   | Benign (SR)                              | Malignant (14.5%)                          | Cystadenofibroma       | -                               |
| 9   | Benign (EA)                              | Malignant (40.3%)                          | Fibroma                | -                               |
| 10  | Benign (SR)                              | Malignant (24.3%)                          | Serous cystadenoma     | -                               |
| 11  | Benign (EA)                              | Malignant (61.2%)                          | Serous cystadenoma     | -                               |
| 12  | Benign (SR)                              | Malignant (26%)                            | -                      | Yes (Benign)                    |
| 13  | Benign (EA)                              | Malignant (35.5%)                          | Hydrosalpinx           | -                               |

SR: Simple Rules (Second step); EA: Expert Assessment (Third step)
Comparison of IOTA three-step strategy and logistic regression model LR2

patients at their initial evaluation. We are aware that the main limitation of our study was the number of patients included, as a larger sample size would have strengthened our results.

Nevertheless, our results for the three-step strategy can be compared to those published about other diagnostic models. For example, the two most commonly used systems such as the Risk of Malignancy Index and the LR2 have shown in their validation studies sensitivity and specificity of 76-87% and 57-97% [32-35] and 92-94% and 75-85%, respectively [13,22,23,30]. Therefore, these results and those obtained in comparison with LR2 in our study indicate a lower diagnostic performance of the rest of the predictive models compared to the three-step strategy. This fact could be validated with direct comparison analysis with a larger sample.

We are aware that the main limitation of our study is the limited number of patients included. We consider that our results can be valid with the sample of 100 patients analysed, although it would be worthwhile corroborating our results can be valid with the sample of 100 patients included. Another limitation of our study may have been that it was carried out in a single hospital, in our case a primary-level hospital without a specialized ovarian cancer unit or expert sonographers, so the inconclusive cases were evaluated by an expert sonographer by means of the electronic submission of ultrasound images, being more appropriate for the expert to evaluate the patients personally. Therefore, the best approach for future studies would be to include a greater number of centers at different levels of care, which would also allow the sample of patients included to be expanded.

In conclusion, our results show that the three-step strategy has high diagnostic accuracy for the classification of benign or malignant adnexal lesions, surpassing the LR2 logistic regression model. These findings suggest that this strategy should be chosen in initial assessments of ovarian pathology by non-expert sonographers, as it also employs easily identifiable ultrasound parameters by most gynaecologists.

Conflict of interest: none

References

1. Cho KR, Shih IM. Ovarian cancer. Annu Rev Pathol 2009;4:287-313.
2. Giede KC, Kieser K, Dodge J, Rosen B. Who should operate on patients with ovarian cancer? An evidence-based review. Gynecol Oncol 2005;99:447-461.
3. Paulsen T, Kjaerheim K, Kaern J, Tretli S, Tropé C. Improved short-term survival for advanced ovarian, tubal, and peritoneal cancer patients operated at teaching hospitals. Int J Gynecol Cancer 2006;16 Suppl 1:11-17.
4. Earle CC, Schrag D, Neville BA, et al. Effect of surgeon specialty on processes of care and outcomes for ovarian cancer patients. J Natl Cancer Inst 2006;98:172-180.
5. Engelen MJ, Kos HE, Willems PH, et al. Surgery by consultant gynecologic oncologists improves survival in patients with ovarian carcinoma. Cancer 2006;106:589-598.
6. Vernooij F, Heintz P, Witteveen E, van der Graaf Y. The outcomes of ovarian cancer treatment are better when provided by gynecologic oncologists and in specialized hospitals: A systematic review. Gynecol Oncol 2007;105:801-812.
7. Timmerman D. The use of mathematical models to evaluate pelvic masses: can they beat an expert operator? Best Pract Res Clin Obstet Gynaecol 2004;18:91-104.
8. Valentijn L, Jurkovic D, Van Calster B, et al. Adding a single CA 125 measurement to ultrasound imaging performed by an experienced examiner does not improve preoperative discrimination between benign and malignant adnexal masses. Ultrasound Obstet Gynecol 2009;34:345-354.
9. Timmerman D, Schwärzler P, Collins WP, et al. Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. Ultrasound Obstet Gynecol 1999;13:11-16.
10. Sokalska A, Timmerman D, Testa AC, Van Holsbeke C, Lissoni AA, Leone FPG. Diagnostic accuracy of transvaginal ultrasound examination for assigning a specific diagnosis to adnexal masses. Ultrasound Obstet Gynecol 2009;34:462-470.
11. Valentijn L, Hagen B, Tingulstad S, Eik-Nes S. Comparison of ‘pattern recognition’ and logistic regression models for discrimination between benign and malignant pelvic masses: a prospective cross-validation. Ultrasound Obstet Gynecol 2001;18:357-365.
12. Lee TS, Kim JW, Park NH, Song YS, Kang SB, Lee HP. Assessing clinical performance of gynaecology residents: sonographic evaluation of adnexal masses based on morphological scoring systems. Ultrasound Obstet Gynecol 2005;26:776-779.
13. Nunes N, Yazbek J, Ambler G, Hoo W, Naftalin J, Jurkovic D. Prospective evaluation of the IOTA Logistic Regression Model (LR2) for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol 2012;40:355-359.
14. Geomini P, Kruijtman WA, Bremer GL, Cnossen J, Mol BW. The accuracy of risk scores in predicting ovarian Malignancy: a systematic review. Obstet Gynecol 2009;113:384-394.
15. Kaijser J, Saysenah A, Van Hoorde K, et al. Presurgical diagnosis using mathematical models and scoring systems: a systematic review and meta-analysis. Hum Reprod Update 2014;20:449-462.
16. Ameye L, Timmerman D, Valentijn L, et al. Clinically oriented three-step strategy for assessment of adnexal pathology. Ultrasound Obstet Gynecol 2012;40:582-591.
17. Saysenah A, Kaijser J, Preisler J, et al. A multicenter prospective external validation of the diagnostic performance of IOTA simple descriptors and rules to characterize ovarian masses. Gynecol Oncol 2013;130:140-146.
18. Testa A, Kajjser J, Wynats L, et al. Strategies to diagnose ovarian cancer: new evidence from phase 3 of the multicentre international IOTA study. Br J Cancer 2014;111:680-688.
19. Peces Rama A, Llanos Llanos MC, Sánchez Ferrer ML, Alcázar Zambrano JL, Martínez Mendoza A, Nieto Díaz A. Simple descriptors and simple rules of the International Ovarian Tumor Analysis (IOTA) Group: a prospective study of combined use for the description of adnexal masses. Eur J Obstet Gynecol Reprod Biol 2015;195:7-11.
20. Alcázar JL, Pascual MA, Graupera B, et al. External validation of IOTA simple descriptors and simple rules for classifying adnexal masses. Ultrasound Obstet Gynecol 2016;48:397-402.
21. Hidalgo JJ, Ros F, Aubá M, et al. Prospective external validation of IOTA three-step strategy to characterize and classify adnexal masses and retrospective assessment of an alternative two-step strategy using simple rules risk. Ultrasound Obstet Gynecol 2019;53:693-700.
22. 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;23:8794-8801.
23. Van Holsbeke C, Van Calster B, Bourne T, et al. External validation of diagnostic models to estimate the risk of Malignancy in adnexal masses. Clin Cancer Res 2012;18:815-825.
24. 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;27:79-105.
25. Timmerman D, Valentin L, Bourne T, et al., International Ovarian Tumor Analysis (IOTA) Group. Terms, definitions and measurements to describe the ultrasonographic features of adnexal tumours: a consensus opinion from the international ovarian tumour analysis (IOTA) group. Ultrasound Obstet Gynecol 2000;16:500-505.
26. Timmerman D, Testa AC, Bourne T, et al. Simple ultrasound based-rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol 2008;31:681-690.
27. Deeks JJ, Altman DG. Diagnostics tests 4: likelihood ratios. BMJ 2004;17:329:168-169.
28. Meys EM, Jeeol LS, Achten NM, et al. Estimating risk of Malignancy in adnexal masses: external validation of the ADNEX model and comparison with other frequently used ultrasound models. Ultrasound Obstet Gynecol 2017;49:784-792.
29. Kajjser J, Van Gorp T, Van Hoorde K, et al. A comparison between an ultrasound-based prediction model (LR2) and the Risk of Ovarian Malignancy Algorithm (ROMA) to assess the risk of malignancy in women with an adnexal mass. Gynecol Oncol 2013;129:377-383.
30. 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 2003;108:2448-2454.
31. Terrin N, Schmid CH, Griffith JL, D’Agostino RB, Selker HP. External validity of predictive models: a comparison of logistic regression, classification trees, and neural networks. J Clin Epidemiol 2003;56:721-729.
32. Moore RG, Jabre-Raughley M, Brown AK, et al. Comparison of a novel multiple marker assay versus the Risk of Malignancy Index for the prediction of epithelial ovarian cancer in patients with a pelvic mass. Am J Obstet Gynecol 2010;203:228.e1-e6.
33. Van Gorp T, Veldman J, Van Calster B, et al. Subjective assessment by ultrasound is superior to the Risk of Malignancy index (RMI) or the Risk of Ovarian Malignancy algorithm (ROMA) in discriminating benign from malignant adnexal masses. Eur J Cancer 2012;48:1649-1656.
34. 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;97:922-929.
35. Bailey J, Tailor A, Naik R, et al. Risk of Malignancy index for referral of ovarian cancer cases to a tertiary centre: does it identify the correct cases? Int J Gynecol Cancer 2006;16 Suppl 1:30-34.