A comparison of CA125, HE4, risk ovarian malignancy algorithm (ROMA), and risk malignancy index (RMI) for the classification of ovarian masses

Cristina Anton, Filomena Marino Carvalho, Elci Isabel Oliveira, Gustavo Arantes Rosa Maciel, Edmund Chada Baracat, Jesus Paula Carvalho

Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, Obstetrics and Gynecology, São Paulo/SP, Brazil

OBJECTIVE: Differentiation between benign and malignant ovarian neoplasms is essential for creating a system for patient referrals. Therefore, the contributions of the tumor markers CA125 and human epididymis protein 4 (HE4) as well as the risk ovarian malignancy algorithm (ROMA) and risk malignancy index (RMI) values were considered individually and in combination to evaluate their utility for establishing this type of patient referral system.

METHODS: Patients who had been diagnosed with ovarian masses through imaging analyses (n = 128) were assessed for their expression of the tumor markers CA125 and HE4. The ROMA and RMI values were also determined. The sensitivity and specificity of each parameter were calculated using receiver operating characteristic curves according to the area under the curve (AUC) for each method.

RESULTS: The sensitivities associated with the ability of CA125, HE4, ROMA, or RMI to distinguish between malignant versus benign ovarian masses were 70.4%, 79.6%, 74.1%, and 63%, respectively. Among carcinomas, the sensitivities of CA125, HE4, ROMA (pre- and post-menopausal), and RMI were 93.5%, 87.1%, 80%, 95.2%, and 87.1%, respectively. The most accurate numerical values were obtained with RMI, although the four parameters were shown to be statistically equivalent.

CONCLUSION: There were no differences in accuracy between CA125, HE4, ROMA, and RMI for differentiating between types of ovarian masses. RMI had the lowest sensitivity but was the most numerically accurate method. HE4 demonstrated the best overall sensitivity for the evaluation of malignant ovarian tumors and the differential diagnosis of endometriosis. All of the parameters demonstrated increased sensitivity when tumors with low malignancy potential were considered low-risk, which may be used as an acceptable assessment method for referring patients to reference centers.

KEYWORDS: Tumor Markers; Biological; Ovarian Neoplasms; CA 125 Antigen; Epididymal Secretory Proteins; Risk Assessment; Sensitivity and Specificity.

INTRODUCTION

Ovarian cancer is the seventh most common cancer in women and is not diagnosed before reaching an advanced stage in approximately 70% of all cases. As a consequence, the 5-year survival rate associated with ovarian cancer is less than 30% (1,2). Therefore, both surgical staging and the performance of optimal cytoreduction procedures in reference centers may have a substantial impact on patient survival.

Over 90% of all ovarian masses detected in pre-menopausal women and up to 60% of masses found in post-menopausal women are benign (3). Thus, it is crucial to assess risk for women who present with pelvic masses, as this process should optimize health policies without overburdening reference centers. Currently, the tumor markers CA125 and human epididymis protein 4 (HE4) as well as the risk ovarian malignancy algorithm (ROMA) and risk malignancy index (RMI) values are used as tools for differentiating between low- and high-risk patients with ovarian cancer. In more than 85% of advanced-stage ovarian carcinoma cases, the levels of CA125 have been found to be elevated above the cutoff value for high-risk patients (>35 U/ml). In contrast, the levels of CA125 are elevated in only 50% of early-stage ovarian carcinoma cases (4). Moreover, in post-menopausal women, CA125 values greater than 95 U/ml are associated with a positive
predictive value of 95% (5). HE4 is a recently discovered tumor marker that has been shown to have a sensitivity of 72.9% and a specificity of 95% for differentiating between types of ovarian masses, and these values are higher than those related to the use of CA125 (6). In 2009, Moore et al. (7) proposed that the ROMA value, which takes into account the levels of CA125 and HE4 together with menopausal status, could be used to evaluate ovarian masses using only quantitative and objective parameters. The use of this algorithm in cohorts of pre- and post-menopausal women resulted in a sensitivity of 88.7% and a specificity of 74.7% (7). Almost 20 years prior to the development of the ROMA, Jacobs et al. (8) created the RMI, which takes into account the CA125 value, menopausal status, and ultrasound parameters. RMI values greater than 200 were shown to be associated with a higher risk of malignancy and demonstrated a sensitivity of 85.4% and a specificity of 96.9% (8).

In this study, we aimed to compare these four methods, namely the CA125 and HE4 measurements and the ROMA and RMI values, in regards to their ability to differentiate low- versus high-risk pelvic masses initially suspected to be of ovarian origin.

PATIENTS AND METHODS

This was a prospective study conducted at the Department of Obstetrics and Gynecology of the Universidade de São Paulo and the Instituto do Câncer do Estado de São Paulo (São Paulo, Brazil) between June 2008 and January 2011. The study was approved by the Ethical Committee for Research Projects of the Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (CAPPesq) (protocol 1067/08), and all of the patients provided informed consent.

Study population

A total of 128 patients between the ages of 15 and 90 years were referred to our hospital with pelvic masses, which were thought to be of ovarian origin as diagnosed by ultrasonography, computed tomography (CT), or magnetic resonance (MR). The women included in the study underwent surgery or imaging-guided biopsy when they presented signs of carcinomatosis. The levels of CA125 and HE4 were measured, and the ROMA and RMI scores were calculated as previously described (7,8). The pre-menopausal patients were evaluated outside of the menstrual period. We considered post-menopausal patients to be those over 50 years of age or those over 40 years of age who had not experienced menses for at least one year. The exclusion criteria included pregnancy, peritoneal dialysis, any previous diagnosed disease commonly associated with an increase in CA 125 (such as mesothelioma), and non-ovarian tumors.

CA125 and HE4 assays

On the day of surgery or imaging-guided biopsy, two blood samples (3-5 ml each) were collected before the procedure. One sample was analyzed for the CA125 level using a Cobas® 4000 analyzer series and the 411 module for immunoassays (Roche®, USA). The second sample was centrifuged and stored at -80°C, and the level of HE4 was subsequently determined using an HE4 enzyme immunoassay (EIA) (Fujirebio Diagnostics Inc., Göteborg, Sweden).

ROMA calculations

The ROMA was calculated as described by Moore et al. (7). The predictive index values were calculated as follows: pre-menopausal Predictive Index (PI) = -12.0 + 2.38 LN (HE4) + 0.0626 LN (CA125); and post-menopausal PI = -8.09 + 1.04 LN (HE4) + 0.732 LN (CA125). In addition, the Predicted Probability (PP) was calculated as: PP = exp(PI) / [1+exp(PI)].

RMI Calculations

The RMI was calculated according to the criteria described by Jacobs et al. (8) as follows: RMI = UxMxserum CA125, where U = 0 for an ultrasound score of 0, U = 1 for an ultrasound score of 1, and U = 3 for an ultrasound score of 2-5 and M = 1 for pre-menopausal women and M = 3 for post-menopausal women. If the patient underwent CT or MR prior to ultrasound, the parameters for the sonographic evaluations were identical to those described by Jacobs et al. (8) and Moore et al. (9). Although the RMI parameters had been validated with ultrasound parameters, we used the same ultrasound parameters for CT and RM as those used by Moore et al. (9).

Cutoff values

The cutoff values for CA125 and HE4 were 35 U/ml (as recommended by the manufacturer) and 70 pM [as used by Moore, et al. (9)], respectively. These cutoff values were the same as those used for the validation of the ROMA (7). The ROMA cutoff values for high-risk patients were ≥13.1% and ≥27.7% for pre-menopausal and post-menopausal women, respectively, as suggested by Moore et al. (7). The cutoff RMI value for differentiating between benign versus malignant masses was 200, as proposed by Jacobs et al. (8). We determined the optimal cutoff value for the CA125 and HE4 tumor markers as well as the ROMA and RMI by analyzing the point of greatest accuracy in the receiver operating characteristic (ROC) curves.

Data Analysis

The sensitivity and specificity for CA125, HE4, ROMA, and RMI were calculated. The ROC curves and area under the curve (AUC) values were calculated to compare the accuracy of each method for predicting malignant ovarian masses. The statistical analyses were performed using MedCalc v11.1.1.0 (MedCalc Software, Mariakerke, Belgium) (10). For all of the statistical comparisons, a level of p<0.05 was accepted as statistically significant.

RESULTS

Eight patients were excluded from this study for the following reasons: technical problems were encountered with the samples from five patients, one patient had a subserosal leiomyoma instead of an ovarian mass, one patient was being treated with peritoneal dialysis, and one patient had been diagnosed with mesothelioma.

The characteristics of the studied population, including age, menopausal status, and the mean and median levels of CA125 and HE4 as well as the RMI and ROMA values, are shown in Table 1. The tumors in this cohort were classified into two groups consisting of those with a low- or high-risk for ovarian malignancy. The low-risk tumors included teratomas (n = 16), endometriomas (n = 13), fibromas (n = 10), and RMI values, in regards to their ability to differentiate low- versus high-risk pelvic masses initially suspected to be of ovarian origin.
The high-risk tumors included low-malignant-potential (LMP) serous (n = 6) and mucinous (n = 11) tumors, serous carcinomas (n = 16), endometrioid adenocarcinoma (n = 3), mucinous adenocarcinomas (n = 3), carcinosarcoma (n = 1), metastatic adenocarcinoma of the breast (n = 1), metastatic adenocarcinomas of the gastrointestinal tract (n = 5), metastatic adenocarcinoma of the colon (n = 1), clear cell carcinoma (n = 1), granulosa cell tumor (n = 4), Leydig cell tumor (n = 1), and steroid cell tumor (n = 1). The sensitivities associated with CA125, HE4, ROMA, and RMI for this cohort of cases were 70.4%, 79.6%, 74.1%, and 63%, respectively.

When the LMP tumors were classified as low-risk, the sensitivities for CA125, HE4, ROMA, and RMI increased to 83.8%, 86.5%, 83.8%, and 75.7%, respectively. In addition, the sensitivities associated with the discrimination between primary carcinomas of the ovary for CA125, HE4, ROMA, and RMI were 93.5%, 87.1%, 95.2%, and 87.1%, respectively. Table 2 provides the sensitivity and specificity values established for CA125, HE4, ROMA, and RMI in both pre- and post-menopausal women. The optimal cutoff values for these data were also calculated. Twelve (92.3%) of the 13 patients diagnosed with endometrioma had elevated values of CA125, and only three (23.1%) had increased HE4 values. The ROC curves for CA125, HE4, ROMA, and RMI were calculated to compare the accuracy of the four methods. The greatest AUC was associated with the RMI values (0.861), as compared to the ROC values for the ROMA (0.824), HE4 (0.777), and CA125 (0.802). The ROC curves were compared using a pairwise comparison method (10), and no differences were detected between the four methods. In addition, no differences were observed in the ROC curves of CA125 and HE4 as compared to the ROMA and RMI. However, differences were observed between the HE4 and ROMA values (p = 0.03) in the overall assessment (Figure 1) and among post-menopausal women (p = 0.05).

DISCUSSION

Our results indicate that the tumor markers CA125 and HE4 as well as ROMA and RMI values are useful methods

### Table 1 - Characteristics of the studied population.

<table>
<thead>
<tr>
<th>Age</th>
<th>Menopausal Status</th>
<th>CA125 (U/ml Mean (SD))</th>
<th>HE4 (pM Mean (SD))</th>
<th>ROMA (%) Mean (SD)</th>
<th>ROMA (%) Median (Range)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre (N)</td>
<td>Post (N)</td>
<td>Median (Range)</td>
<td>Median (Range)</td>
<td>Median (Range)</td>
</tr>
<tr>
<td>Benign</td>
<td>50.7</td>
<td>51 29</td>
<td>37</td>
<td>36.8 (55.5)</td>
<td>17.1</td>
</tr>
<tr>
<td>LMP*</td>
<td>56.4</td>
<td>58 6</td>
<td>11</td>
<td>147 (266.5)</td>
<td>26.4</td>
</tr>
<tr>
<td>Malignant</td>
<td>54.7</td>
<td>54 12</td>
<td>25</td>
<td>318.9 (517.8)</td>
<td>138.7</td>
</tr>
</tbody>
</table>

* LMP: Low Malignant Potential.

### Table 2 - The sensitivity and specificity values associated with CA125, HE4, ROMA, and RMI for pre- and post-menopausal patient groups at the optimal identified cutoff value and the standard cutoff value for each method.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Optimal cutoff</th>
<th>Standard cutoff****</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ROC/AUC</td>
<td>Cutoff Value</td>
</tr>
<tr>
<td>All Patients</td>
<td>CA125</td>
<td>0.802</td>
</tr>
<tr>
<td></td>
<td>HE4</td>
<td>0.777</td>
</tr>
<tr>
<td></td>
<td>ROMA</td>
<td>0.824</td>
</tr>
<tr>
<td></td>
<td>RMI</td>
<td>0.861</td>
</tr>
<tr>
<td>Pre-menopausal</td>
<td>CA125</td>
<td>0.807</td>
</tr>
<tr>
<td></td>
<td>HE4</td>
<td>0.774</td>
</tr>
<tr>
<td></td>
<td>ROMA</td>
<td>0.791</td>
</tr>
<tr>
<td></td>
<td>RMI</td>
<td>0.875</td>
</tr>
<tr>
<td>Post-menopausal</td>
<td>CA125</td>
<td>0.846</td>
</tr>
<tr>
<td></td>
<td>HE4</td>
<td>0.762</td>
</tr>
<tr>
<td></td>
<td>ROMA</td>
<td>0.840</td>
</tr>
<tr>
<td></td>
<td>RMI</td>
<td>0.853</td>
</tr>
</tbody>
</table>

*ROC/AUC = area under the ROC (Receiver Operating Characteristic) curve.
**Sen = Sensitivity.
***Spe = Specificity.
****Standard cutoff: CA125, Bast et al., 1983; HE4, Moore et al., 2008; ROMA, Moore et al. 2010; RMI, Jacobs et al., 1990.
*****Cutoff values used for ROMA are different for pre- and post-menopausal women.
for differentiating ovarian masses according to whether they are associated with a high or low risk for developing into ovarian cancer, and this type of assessment which will ultimately optimize the referral of patients to reference centers. We intentionally modified the above methods to perform an analysis using the available tools in routine practice. For the ROMA, we used the CA125 kit from Roche® even though this algorithm had been validated using the Fujiribio CA125 kit, as we routinely use the Roche® kit for our services. Although we wanted to ensure that the current study provided an accurate representation of the tools available at our hospital, we applied the same numeric cutoff values proposed by Moore et al. (7). Moreover, we believe that the small deviations observed in the analysis presented here did not compromise the overall findings of the study.

The RMI method originally described by Jacobs et al. (8) used ultrasound assessment as the sole imaging parameter, although we included both CT and MRI in our study (as did Moore et al.) (9). It would be expected that this modification would improve the results for the RMI than those presented by Jacobs et al., as CT and MR technologies can more accurately assess adnexal tumors. However, our study found that this modification did not provide more accurate results. In the study by Jacobs et al. (8), the sensitivity and specificity reached 85.4% and 96.9%, respectively, using a cutoff value of 200. In the present study, the sensitivity and specificity values associated with the RMI were 66.7% and 87.9%, respectively, which are lower than those demonstrated by Jacobs et al. (8). One potential explanation for these different results may have been the greater heterogeneity of histologic types observed in our study. However, the samples we analyzed demonstrated a profile very similar to that observed in clinical practice.

To evaluate the accuracy of the four ovarian cancer risk stratification methods, an AUC was calculated for each ROC. We calculated ROC values of 80.2% and 77.7% for CA125 and HE4, respectively, and found no statistical difference between these values \( p = 0.67 \). In contrast, an evaluation of the same tumor markers by Moore et al. (9) found ROC values of 83.6% and 90.8%, respectively, and no statistically significant differences were observed for these ROMA and RMI curves either \( p = 0.50 \). However, the study by Moore et al. (9) reported the AUCs for the ROMA and RMI to be 91.3% and 84.4%, as compared to the values of 82.4% and 85.5% obtained in our study, respectively. These differences may also be explained by the more heterogeneous quality of the tumors evaluated in our study.

In the current study, the optimal cutoff values associated with the ROMA were 13.97% and 39.68% for pre- and post-menopausal women, respectively. The pre-menopausal ROMA cutoff value in our study was similar that reported by Moore et al. (9), whereas the post-menopausal value was higher (13.1% and 27.7%, respectively). The cutoff values reported by Van Gorp et al. (11) were 16.6% and 35.9% for pre- and post-menopausal women, respectively, which were also similar to those identified in our study. The authors of this previous study also found an overall cutoff value for the post-menopausal ROMA of 22.2%, which was similar to the value of 23.3% identified in our study. The use of different CA125 kits may have resulted in these differences, and the inclusion of CT and MRI methods may explain the higher cutoff RMI value obtained in the current study as compared to the study by Jacobs et al. (8) (275.7 vs. 200, respectively).

One major problem associated with the pre-operative evaluation of pelvic masses concerns the identification of LMP tumors and the diagnosis of endometriosis. We found that the four methods of risk evaluation demonstrated better performance when LMP tumors were classified as low-risk tumors. The sensitivity value for CA125 detection was 83.8% with a specificity of 71.1%, whereas these values were 70.4% and 74.2%, respectively, when the tumors were classified as high-risk. In addition, HE4 sensitivity increased from 79.65 to 86.5%, ROMA sensitivity increased from 74.1% to 83.8%, and RMI sensitivity increased from 63% to 75.7% when these tumors were classified as low-risk. Although the classification of LMP tumors remains controversial, clinical and biological evidence suggests that these tumors can be classified as low-risk. LMP tumors are associated with a good prognosis, and 5-year survival rates are approximately 98% for stage I tumors and 90% for stage III tumors with non-invasive implants (12). In addition, BRAF and KRAS mutations characterize the low-grade pathway of ovarian carcinogenesis, and borderline tumors are identified as precursor lesions of this pathway (13). Although it is outside the scope of the current study to discuss the nature and treatment options for LMP tumors, these types of tumors represent an important differential diagnosis for pelvic masses. In our cohort, these tumors were found in 17 (14.2%) cases, which mirrors the frequency of these tumors observed in common clinical practice. Because these tumors are low-grade and in most cases do not require surgical biopsy as a first-line approach, their inclusion in the low-risk category appears to be acceptable. Therefore, patients receiving care from non-specialized centers may also be referred to reference centers for appropriate staging without compromising their prognosis (12). Moreover, if LMP tumors are classified as low-risk, the burden related to an influx of lower-risk patients at specialized centers may be reduced.

The inclusion or exclusion of LMP and other non-epithelial tumors in the assessment of data is one of the most important factors affecting the reported accuracy of these methods. In the study by Moore et al. (6), the sensitivity of HE4 was found to be 72.9% at a 95% specificity. In our analysis, only epithelial ovarian cancers were included, and therefore nine LMP tumors, two non-epithelial ovarian tumors, and 13 metastatic tumors of the ovary were excluded. In 2011 from an analysis of patients with ovarian cancer, Chang et al. (14) evaluated 491 patients and obtained a sensitivity of 73% and 88% using the markers HE4 and CA125, respectively. As a result, the sensitivity value for HE4 was found to be 79.6%, which was higher than that reported by Moore et al. (6). As previously...
stated, the sensitivity of HE4 in our study was determined to be 86.5% with the exclusion of LMP tumors. The sensitivity and specificity associated with the ROMA values for the cases analyzed were 74.1% and 75.8%, respectively. These results were not consistent with those of Moore et al. (6), which could have been due to the exclusion of non-epithelial tumors in the previous analysis (7).

Another important differential diagnosis for pelvic masses is endometriosis, which is the disease that most often interferes with the accuracy of the methods used for pre-operative evaluations of cancer risk. In the present study, 12/13 (92.3%) patients with endometriosis had CA125 values above 35 U/ml, whereas only 3/13 (23.1%) patients had HE4 values above 70 pM. The increased specificity of HE4 for the differentiation between endometriosis and ovarian cancer is in agreement with two recently published studies (15,16), suggesting that the use of both markers together can improve this type of evaluation.

Despite small variations, the four methods that were evaluated for their ability to differentiate adnexal masses (CA125, HE4, ROMA, and RMI) demonstrated similar levels of accuracy. The RMI was found to have the lowest sensitivity but provided the best numeric accuracy of the four methods. The tumor marker HE4 demonstrated the best overall sensitivity for the evaluation of malignant ovarian tumors and the differential diagnosis of endometriosis. All of the parameters demonstrated increased sensitivity when tumors with low malignancy potential were considered low-risk, which may be used as an acceptable assessment method for referring patients to reference centers.

**AUTHOR CONTRIBUTIONS**

Anton C conceived the study, collected the data, participated in the samples analysis, drafted the manuscript and performed the statistical analysis. Carvalho FM conceived the study and participated in its design, coordination and drafting. Oliveira EI participated in the samples analysis and interpretation. Maciel GA participated in the sample analysis. Baracat EC reviewed the manuscript. Carvalho JP conceived the study and participated in its design, coordination and drafting. All authors have read and approved the final manuscript.

**REFERENCES**